

UNCLASSIFIED

AD NUMBER
AD222927
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; JUN 1955. Other requests shall be referred to Army Ordnance Corps, Office of Research, Durham, NC.
AUTHORITY
usaro ltr, 15 mar 1971

THIS PAGE IS UNCLASSIFIED

UNCLASSIFIED

AD 222 927

*Reproduced
by the*

**ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA**



UNCLASSIFIED

**Best
Available
Copy**

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

UNANNOUNCED

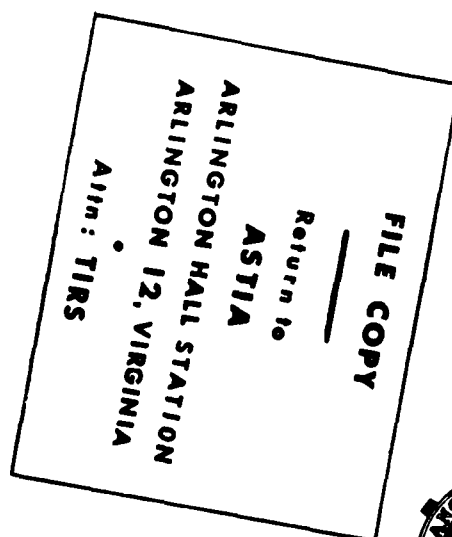
AD No. 222927

ASTIA FILE COPY

Office of Ordnance Research

OORR 55-2

**Transactions of The First Conference
of Arsenal Mathematicians**



OFFICE OF ORDNANCE RESEARCH
ORDNANCE CORPS, U. S. ARMY
BOX CM, DUKE STATION
DURHAM, NORTH CAROLINA

OFFICE OF ORDNANCE RESEARCH

Report No. 55-2

June 1955

**TRANSACTIONS OF THE FIRST CONFERENCE
OF ARSENAL MATHEMATICIANS**

**held at Watertown Arsenal
October 29, 1954**

**FOREIGN ANNOUNCEMENT AND DISSEMINATION
OF THIS REPORT BY ASTIA IS LIMITED**

**Office of Ordnance Research
Ordnance Corps, U. S. Army
Box OM, Duke Station
Durham, North Carolina**

FOREWORD

The First Conference of Arsenal Mathematicians was held 29 October 1954 with Watertown Arsenal acting as host. Colonel A. P. Taber, Commanding Officer at Watertown Arsenal, said words of greeting to the fifty-six members of the Conference. Colonel P. N. Gillon, Ordnance Corps, Commanding Officer of the Office of Ordnance Research, served as chairman of the meeting. Colonels Gillon and Taber pointed to specific ordnance projects needing a high level of scientific knowledge, adding that applied mathematics plays a vital role in many of these, and advancements in this field are needed if high level improvements of ordnance needs are to continue.

With the one exception of the invited address by Professor F. J. Murray of Columbia University, the talks were given by full-time employees of six different government installations. All the papers covered points of common interests to those in attendance and were of such high mathematical caliber that a general request was made to have the manuscripts bound in a single research report. In addition to affording an opportunity to the speakers to present their new results, the conference permitted those in one arsenal to see and discuss the problems faced by those in another arsenal. Such interplay of ideas followed by small group discussions is vital to continued new attacks on the many problems facing arsenal mathematicians.

TABLE OF CONTENTS

	Page
Foreword	iii
Program	vii
Stress Analysis in Small Arms Barrels By A. Hammer	1
A Note on Zero Crossings and Average Frequency By P. R. Karr	2
Influence of Elastic Deformation on Accuracy of High Performance Rockets By H. J. Stewart	9
The Use of Analog Computers in the Design of Military Vehicles By Thomas Wood	10
Problems in the Application of Least Squares to Function Generator Design By J. C. Tappert	33
Some Typical Problems Confronting Mathematicians at BRL By Saul Gorn	39
Computing Methods for the Solution of Systems of Differential Equations By F. J. Murray	52

Initial Distribution

The initial distribution list of the Transactions of the First Conference of Arsenal Mathematicians includes those who attended the meeting and/or the government installations with which they are associated. For economy, only a limited number of copies have been sent to each. Additional copies will be transmitted upon request.

PROGRAM

CONFERENCE OF ARSENAL MATHEMATICIANS

WATERTOWN ARSENAL

29 October 1954

**Chairman: Colonel P. N. Gillon, Ordnance Corps, Commanding Officer
of the Office of Ordnance Research**

MORNING SESSION 0900 - 1150

Introductory Remarks

**Stress Analysis in Small Arms Barrels (30 minutes)
Dr. A. Hammer, Springfield Arsenal**

**Theory of Noise Modulation (30 minutes)
Dr. P. R. Karr, Diamond Ordnance Fuse Laboratories**

Break

**Influence of Elastic Deformation on Accuracy of High Performance
Rockets (30 minutes)
Dr. H. J. Stewart, Jet Propulsion Laboratory**

**The Application of Electronic Analog Computers to Problems
of Tank Design (15 minutes)
Mr. T. A. Wood, Detroit Arsenal**

Lunch 1150 - 1250

Tour of Watertown Arsenal 1250 - 1330

AFTERNOON SESSION 1330 - 1615

**Problems in the Application of Least Squares to Function
Generator Design (30 minutes)
Dr. J. C. Tappert, Frankford Arsenal**

**Some Typical Problems Confronting Mathematicians at the
Ballistic Research Laboratories (30 minutes)
Dr. S. Gorn, Aberdeen Proving Ground**

Break

**Computation Study of Differential Equations (50 minutes)
Professor F. J. Murray, Columbia University**

THE FIRST CONFERENCE
OF
ARSENAL MATHEMATICIANS

Watertown Arsenal

29 October 1954

Colonel P. N. Gillon, Chairman

STRESS ANALYSIS IN SMALL ARMS BARRELS
By A. Hammer

Increased firing rates and the ever increasing severity of the firing schedules induce conditions in barrels of small arms weapons, from Caliber .30 to 30mm, which require different design and analytical approaches from those used in the past. Vital points of new barrel constructions are described and discussed and brief outlines of the processes which occur within lined barrels are given in the first part of the paper. The barrel stress analysis, which is the subject of the second part, considers the thermal, shrink-fit and rifling stresses as well as the pressure stresses, and the graphical presentation of the total equivalent stresses facilitates the task of establishing optimum design parameters for lightweight accurate barrels of extended life.

- - - - -

The main portion of the address by Dr. A. Hammer can be found in the security information paper entitled "Barrel Research, Development and Stress Analysis at Springfield Armory." Copies of this paper may be obtained from Springfield Armory by authorized persons.

- - - - -

A NOTE ON ZERO CROSSINGS AND AVERAGE FREQUENCY

By P. R. Karr

1. Introduction

The function $\sin(2\pi g_0 t + \delta)$ with the constant frequency g_0 , the constant angle δ , and t denoting the time variable, becomes zero $2g_0$ times per unit time. A natural question that arises is: if in the above expression the phase function $2\pi g_0 t$ is replaced by a random function of time, at what average rate will zero-crossings occur and to what extent can this rate be associated with an "average frequency" of the random function? This question is discussed in the following.

2. Development

We consider first the function $\sin(N(t) + \delta)$ where δ is a constant as above and $N(t)$ is a stationary random ("noise") function of t ; in the present paper we assume that $N(t)$ has a Gaussian amplitude distribution. To find the average number, n , of zeros of $\sin(N(t) + \delta)$ we first find the average number of times per unit time that the argument $N(t) + \delta$ crosses the values $0, \pm\pi, \pm2\pi$, etc; we then add all these results together, giving an infinite series, whose sum, it is seen, is the desired result, n .

(1)

The fundamental paper of Rice treats the problem of the rate of zero-crossings of $N(t)$. To make the calculations listed in the preceding paragraph, we should like a formula for the expected number of "A-crossings", that is, the expected number of crossings per unit time of the arbitrary value "A" by the noise function $N(t)$. It turns out that such a formula can be obtained by a simple extension of Rice's derivation for the zero-crossings. The result is

$$r = \left[\frac{-\ddot{R}(0)}{\pi^2 R(0)} \right]^{1/2} \exp \left[-\frac{A^2}{2R(0)} \right] \quad (1)$$

where

r is the expected number per unit time of crossings of the value "A" by $N(t)$,
 $R(t)$ is the correlation function of $N(t)$,
 $\ddot{R}(t)$ is the second derivative of $R(t)$ with respect to t ,

and $R(0)$ and $\ddot{R}(0)$ are the values of $R(t)$ and $\ddot{R}(t)$ at $t = 0$. As is well known, $R(0)$ is the mean intensity of $N(t)$, or $\overline{N^2}$.

Now $N(t) + \delta$ has the value of $A + \delta$ when $N(t)$ has the value of A , so to obtain n we need to give A in eg. (1) the values $k\omega - \delta$ for all integral values of k , and add the resulting terms. We have then

$$n = \left[\frac{-R(0)}{n^2 R(0)} \right]^{1/2} \sum_{k=-\infty}^{\infty} \exp - \frac{(k\omega - \delta)^2}{2R(0)} \quad (2)$$

Eq. (2) is a formula, to be discussed in more detail below, giving n in terms of δ , $R(0)$ and $\ddot{R}(0)$. Before considering the dependence upon these factors however, we wish to transform (2) into a form which will exhibit specifically the "average frequency" of $N(t)$.

We have, by the Wiener-Kinchine theorem ⁽²⁾,

$$R(t) = \int_0^{\infty} S_N(f) \cos 2\pi f t \, df \quad (3)$$

where S_N is the spectral intensity or "power spectrum" of $N(t)$ and f denotes frequency.

Then

$$-\ddot{R}(t) = \int_0^{\infty} 4\pi^2 f^2 S_N(f) \cos 2\pi f t \, df. \quad (4)$$

Also

$$R(0) = \int_0^{\infty} S_N(f) \, df \quad (5)$$

and

$$-\ddot{R}(0) = \int_0^{\infty} 4\pi^2 f^2 S_N(f) \, df. \quad (6)$$

Now we may introduce the "instantaneous frequency" of $N(t)$ by assuming it to be expressible in the form

$$N(t) = 2\pi \int_{-\infty}^t g(t) dt \quad (7)$$

Eq. (7) expresses the relation between $N(t)$ and $g(t)$ as that between phase angle and instantaneous frequency. Since the transfer function of a perfect integrator is $\frac{1}{2\pi jf}$, we have

$$S_N(f) = \frac{4\pi^2}{4\pi^2 f^2} S_g(f) = \frac{1}{f^2} S_g(f) \quad (8)$$

where $S_g(f)$ is the spectral intensity of $g(t)$.

Therefore

$$- \ddot{R}(0) = \int_0^{\infty} 4\pi^2 S_g(f) df. \quad (9)$$

Now it is a well known property of stationary random processes such as $g(t)$ that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g^2(t) dt = \int_0^{\infty} S_g(f) df = \overline{g^2} \quad (10)$$

where $\overline{g^2}$ denotes the time average (or stochastic average) of g^2 .

Eq. (9) therefore has the interpretation that if a noise function

$N(t)$ may be expressed as the integral of an instantaneous frequency, and if $R(t)$ is the correlation function of $N(t)$, then $-R(0)$ is $4\pi^2$ times the mean square frequency g^2 of $N(t)$.

At this point we introduce the average value of the absolute value $\overline{|g|}$. (The average \overline{g} is zero.) If $g(t)$ has a Gaussian amplitude distribution, which we assume, and which is consistent with the previous assumption that $N(t)$ is Gaussianly distributed, then it is easily found that

$$\overline{|g|} = \sqrt{\frac{2}{\pi}} [\overline{g^2}]^{1/2}. \quad (11)$$

Introducing the last two results we may now rewrite (2) in the form

$$n = \sqrt{\frac{2\pi}{R(0)}} \overline{|g|} \sum_{k=-\infty}^{\infty} \exp \frac{(-kn - \gamma)^2}{2R(0)} \quad (12)$$

$$= \sqrt{\frac{2\pi}{R(0)}} \overline{|g|} \exp \frac{-\gamma^2}{2R(0)} \left\{ \frac{1}{2} + \sum_{n=1}^{\infty} \exp \frac{-n^2 \pi^2}{2R(0)} \right\} \cosh \frac{n\pi\gamma}{R(0)} \quad (13)$$

An application of Poisson's Formula⁽³⁾ to (13) gives the alternate form

$$n = 4\overline{|g|} \left\{ \frac{1}{2} + \sum_{n=1}^{\infty} \exp \frac{-2R(0)n^2}{\pi^2} \cos 2\gamma n \right\} \quad (14)$$

The task of numerical computation is greatly facilitated by these two alternate forms since (13) converges rapidly when $R(0)$ is small, while (14) is rapidly convergent when $R(0)$ is large. Thus the two alternate forms complement each other nicely.

From (14) we have the interesting result that for large $R(o)$, and independently of δ , n is asymptotically $2|g|$, that is, twice the average frequency (the cognomen "average frequency" applied to $|g|$ implies of course that negative and positive frequencies are regarded as on an equal basis). For small $R(o)$ n is asymptotically

$$\sqrt{\frac{2\pi}{R(o)}} |g| \left\{ \exp \frac{-\delta^2}{2R(o)} + \exp \frac{-(n-\delta)^2}{2R(o)} \right\}.$$

In Fig. 1 there is plotted the function $\frac{n}{|g|}$ for various values of $R(o)$ and δ . It will be seen that n can be greater than $2|g|$, as well as less. The influence of δ becomes rapidly dissipated as $R(o)$ increases. In fact Fig. 1 shows that even for $R_o = \frac{\pi^2}{4}$, n is already essentially equal to $2|g|$.

One may ask, what is the value of n when averaged over δ ? The form (14) is convenient for examining this question. One sees that all the $\cos 2\delta n$ terms drop out upon averaging and we are left with

$$n_{av} = 2|g| = \left[\frac{-2\ddot{R}(o)}{3} \right]^{1/2} \quad (15)$$

Thus the average value of n is the same as the asymptotic value for high mean intensity $R(o)$, namely twice the average frequency. Thus for either of these conditions the concept "average instantaneous frequency" is useful and valid, since the expected number of zero crossings is twice the average frequency. This is an intuitively satisfying generalization of the corresponding result for $\sin[2\pi \int_0^t g_o dt + \delta]$; that is under the

conditions indicated, if g_o in the preceding expression be replaced by $g(t)$ then the result $n = 2g_o$ is generalised to $n = 2|g|$ or $n_{av} = 2|g|$.

3. Summary

An expression is derived for the expected rate n of zero crossings

of the function $\sin[N(t) + \delta]$, where $N(t)$ is a stationary, Gaussianly distributed, random or "noise" function with zero mean. The expression for n involves $R(0)$, the second derivative with respect to t of the correlation function $R(t)$ of $N(t)$ at $t = 0$, $R(0)$, and δ . When $N(t)$ is expressed as the integral of an instantaneous frequency function $g(t)$, in the form $N(t) = 2\pi \int_{-\infty}^t g(t) dt$, $R(0)$ may be expressed in terms of

$\overline{|g|}$, the average value of $|g|$. The resulting expression for n is proportional to $\overline{|g|}$, the proportionality factor being a function of δ and $R(0)$, the mean intensity of $N(t)$. When $R(0)$ is large, this factor is, closely equal to 2 for any value of δ , while for any value of $R(0)$, the value of this proportionality factor averaged over all values of δ is exactly 2. These results are in pleasing analogy with the case in which $g(t)$ is replaced by the constant g_0 , in which case n is $2g_0$. Thus the concept of "average instantaneous frequency" is a useful one in the sense that the ratio of 2 that holds between n and the fixed frequency g_0 also holds between n and the average frequency $\overline{|g|}$ if either of the above conditions is considered. If $R(0)$ is small and averaging over δ is not allowed, the situation is less simple. Details are given in formulas and in Fig. 1.

4. Acknowledgements

This work was stimulated by discussions with Mr. B. M. Horton who has been interested for some time in the concept of "average frequency".

Most of the numerical work associated with Fig. 1 was performed by Mr. C. Morrison.

-
- (1) S. O. Rice, Bell Telephone System Monograph 1589, p. 57.
 - (2) N. Wiener, Acta Math. 55, 117 (1930).
 - (3) Morse and Feshbach "Methods of Theoretical Physics" p. 467.

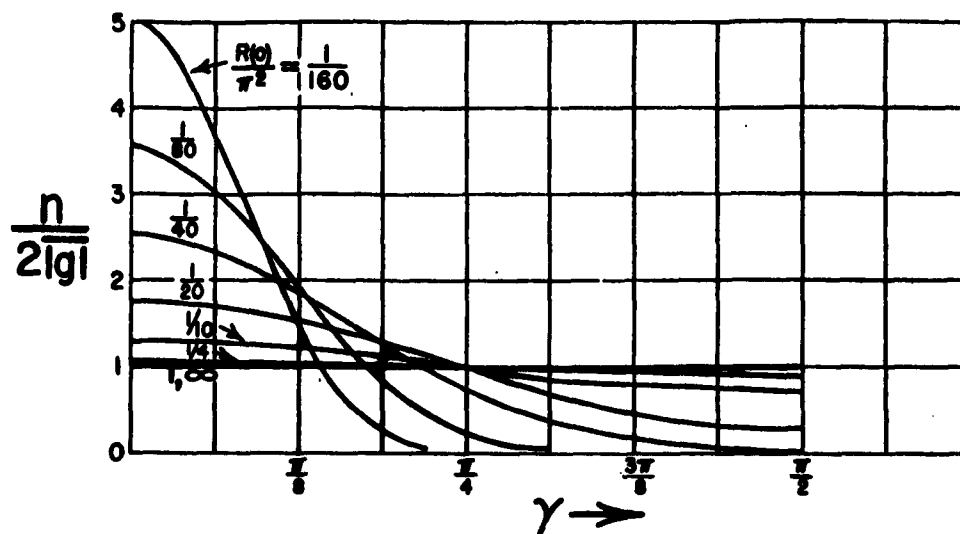


Figure 1a. $\frac{n}{2|q|}$ vs γ for various values of $R(0)$

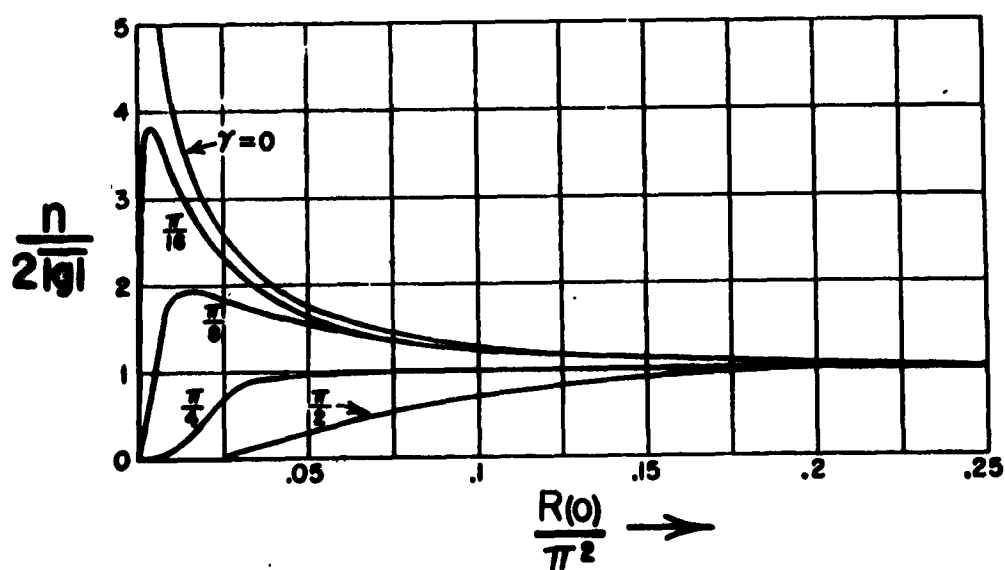


Figure 1b. $\frac{n}{2|q|}$ vs $R(0)$ for various values of γ

INFLUENCE OF ELASTIC DEFORMATION
ON ACCURACY OF HIGH PERFORMANCE ROCKETS

By H. J. Stewart

- - - - -

The paper presented by Dr. H. J. Stewart contained some material abstracted from reports having a Confidential classification. Those interested in pursuing further the subject matter he discussed are referred to the following articles

1. JPL Publication No. 24, by J. E. Froehlich (1953).
2. JPL Progress Report No. 20-250, by I. Statler (In publication).
3. JPL Progress Report No. 20-251, by J. Lorell (In publication).

- - - - -

THE USE OF ANALOG COMPUTERS
IN THE DESIGN OF MILITARY VEHICLES

By Thomas Wood

The development of military tanks, like other automotive vehicles, depends not only on mathematical computation but also on a procedure based on construction of models and extensive tests. It has not been unusual to construct complete vehicles, subject these to severe laboratory and arduous road tests, only to discover that the new design may not have contributed materially to the desired performance characteristics of the vehicle. Add to this a shortage of trained personnel capable of conducting a mathematical analysis and it can be readily understood why costs of engineering development may grow out of proportion.

Thanks to electronics, relatively low cost, high speed analog computers have become available. Engineers at the Detroit Arsenal recognized the possibilities of this instrument and now utilize the computer in the early stages of the development of a complete vehicle or its complex components. The advantages which result from this rapid analytical investigation over the time consuming, tedious and costly manual computations are numerous, but more important is the fact that new ideas and design principles can be investigated almost as rapidly as they are conceived.

The computer makes it possible to economically determine the feasibility of a design without constructing expensive models and conducting extensive laboratory tests which are often destructive. Moreover, it helps the engineer to determine the design principles which offer the most promising avenues of investigation. In cases where uncertainty exists concerning the values of parameters, the computer aids in defining the limits within which the parameters can vary to produce satisfactory results. This feature is not only advantageous in design and development, but is also of extreme value in establishing tolerances on vehicle parts which have been approved for production. This prevents the costly manufacture of precision parts for a system in which precision is not necessary for performance.

An analog computer, as its name implies, is an electronic instrument capable of making high speed mathematical calculations, once the physical behavior of a device has been described and formulated into mathematical statements. This can be done because the behavior of most physical systems can usually be expressed in terms of mathematical equations and graphs. When the mathematical or graphical statements become available and the various elements (multipliers, integrators, etc.) of the computer are connected in a manner analogous to the equations, the computer will perform the appropriate mathematical operations.

The more frequent fundamental operations performed by analog computers are:

1. Multiplication by a constant.
2. Integration of quantities with respect to time.
3. Multiplication of variables.

The operational amplifier is the basic building block of the computer. The various operations which it performs depends on the impedances that are placed in the feedback and input circuits of the amplifier. All the physical parameters are expressed in terms of voltages.

To illustrate the method by which the operational amplifier of the computer functions as a multiplier, let us assume that (Figure 1) the gain G of the amplifier approaches infinity.

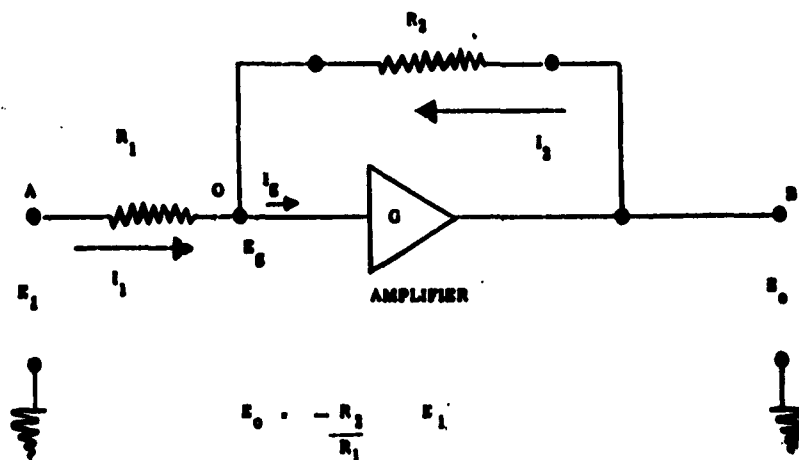


FIGURE 1 - MULTIPLICATION BY A CONSTANT

To get a finite voltage out of the amplifier at B we must not have any voltage at point O; hence E_g will be zero. In a practical case we may use a gain in the order of 200,000 which makes E_g not zero but of such magnitude that it can be neglected. By placing resistors R_2 and R_1 in the feedback and input circuits respectively, the amplifier is said to be "patched" for multiplication. As an example, let us set the operational amplifier to multiply by two by placing a one megohm resistor in the feedback, and a .5 megohm resistor in the input. The operation is performed because any voltage E_1 impressed at A across the input resistor R_1 causes a current I_1 to flow in the input circuit. A fundamental electrical theorem (Kirchhoff's law) states that the algebraic sum of the currents that meet at any junction is zero. Therefore $I_1 + I_2 + I_g$ must equal to zero. I_g is the current drawn by the grid of the operational amplifier and is on the order of 10^{-10} amperes, therefore, it can be neglected. Thus I_2 must be equal in magnitude and opposite in sign to I_1 to make the sum equal to zero. This produces a sign change across the amplifier and will induce a voltage E_o which is of opposite sign and twice E_1 . This is true because I_2 is equal to I_1 and R_2 is twice R_1 . Hence, E_o is twice E_1 . To multiply by any constant, it is necessary to select the appropriate R_1 . (R_2 is usually kept at 1 megohm).

Figure 2 illustrates the method of "patching" the operational amplifier for use as an integrator.

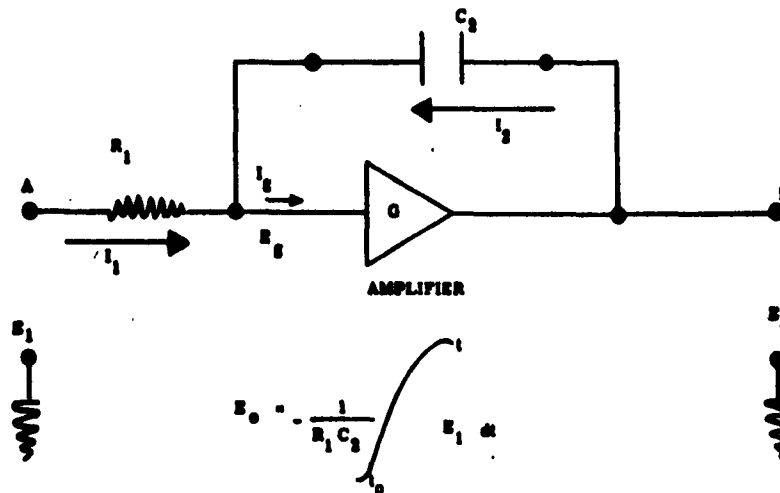


FIGURE 2 - INTEGRATION

By placing a 1 megohm resistor in the input circuit (R_1) and a one microfarad capacitor (C_2) in the feedback circuit, the operational amplifier will generate the time integral of an input voltage E_1 . Again using Kirchoff's law we see that the voltage at A, called E_1 , produces a current I_1 to flow through resistor R_1 . In order to satisfy the law that the algebraic sum of the currents which meet at any junction is equal to zero, I_2 must be equal in magnitude and opposite in sign to I_1 . To produce this current I_2 the voltage E_2 across C_2 must be increased at a constant rate; hence E_2 is

the integral of E_1 .

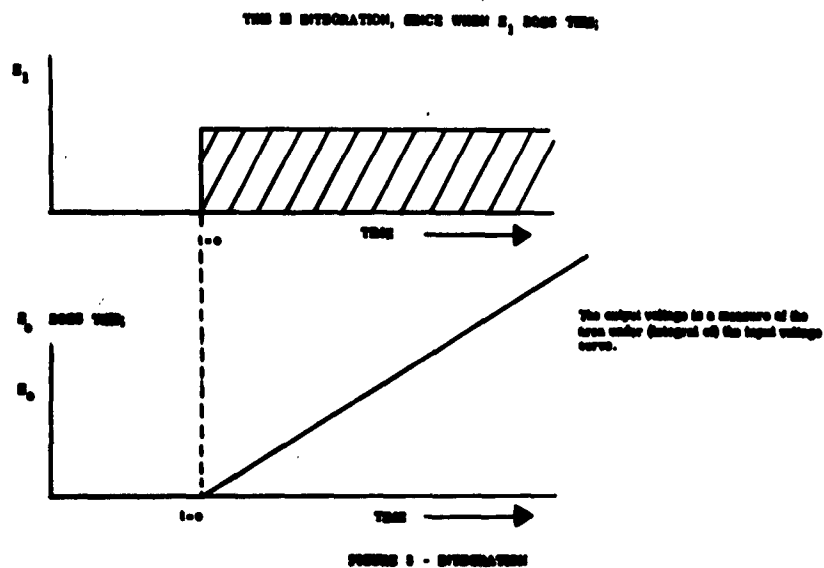


Figure 3 shows that the slope of E_0 is dependent upon the ratio of C_2 and R_1 . Therefore, we can both integrate and multiply by a constant with this arrangement.

Multiplication of two variables can be performed by means of a servo system.

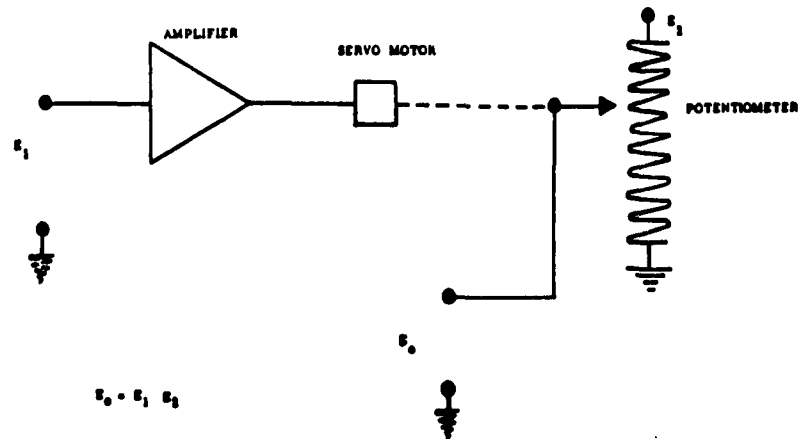


FIGURE 4 - MULTIPLICATION OF TWO VARIABLES

In Figure 4 the variable E_1 produces an output shaft motion whose angular position varies with respect to time exactly as E_1 . The output shaft controls the position of the arm of a potentiometer across which the variable E_2 is applied. The voltage on the arm is a fractional part of the applied voltage E_2 , or $E_0 = E_1 E_2$.

Although the possibility of using an analog computer for design and development of military vehicles was apparent, there was no way of determining the accuracy inherent in this particular type of simulation. For this reason, a research program was instituted at the Willow Run Research Center, under contract with the Detroit Arsenal. The

purpose of the research was to correlate the response of an actual tank suspension system with that of a simulated tank.

An M47 tank, weighing approximately 48 tons, was selected for experiments in order to correlate the simulated conditions with the experimental data. The features of the tank are shown in Figure 5.

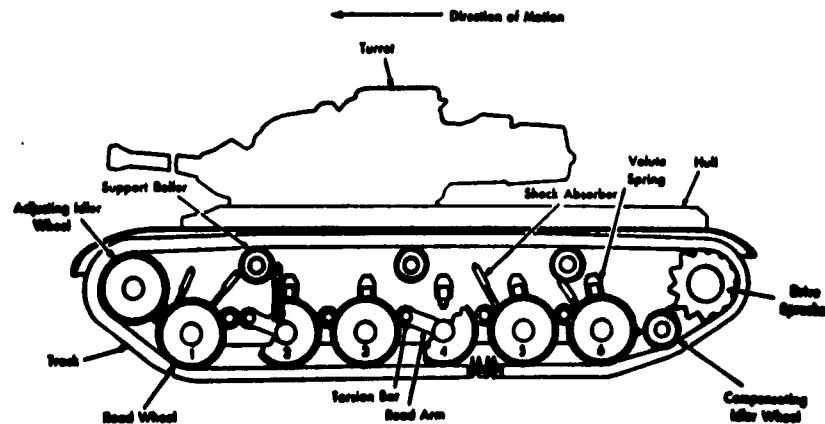


Fig. 5 FEATURES OF M47 SUSPENSION SYSTEM

The hull is supported by twelve road wheels, six on each side. These wheels are attached to the hull through road arms,

shock absorbers, and torsion bars which in effect act as springs. Volute springs, or bump stops, as they are commonly called, are installed to limit the deflection of the wheel assembly with respect to the hull. Attached to road wheels 1, 2, 5 and 6 are shock absorbers whose function is to damp out motion of the hull. To increase the mobility of the tank over rough terrain, a track is provided on each side of the tank.

The field tests consisted of running the tank at various speeds (2 to 20 mph) over two types of road obstacles. One obstacle was a 12" and 12" timber, and the other an inverted V ramp 10' long and 8" high, as shown in Figure 6.

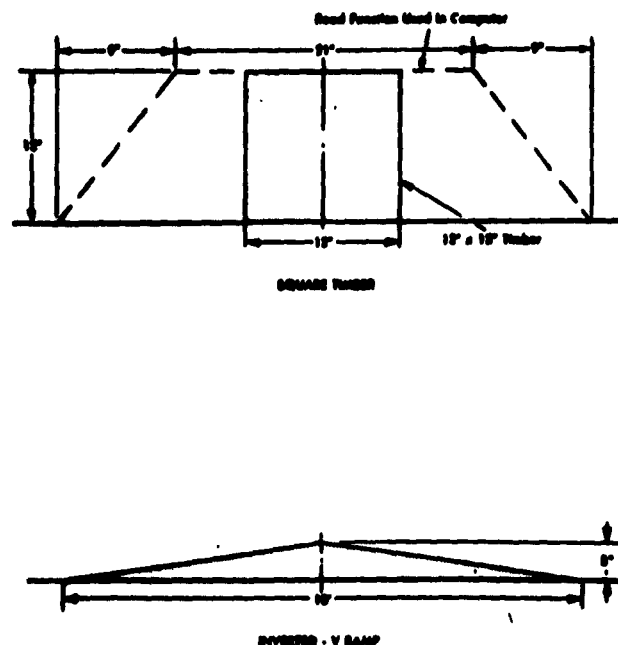


Fig. 6 ROAD FUNCTIONS USED FOR CORRELATION TESTS

NOTES:

1. Positive direction of force and displacement is assumed upward. Positive direction of moment and angle is assumed clockwise.
2. Reference position of hull and each wheel is vertical position of each item if tank were resting on level ground and had no weight.

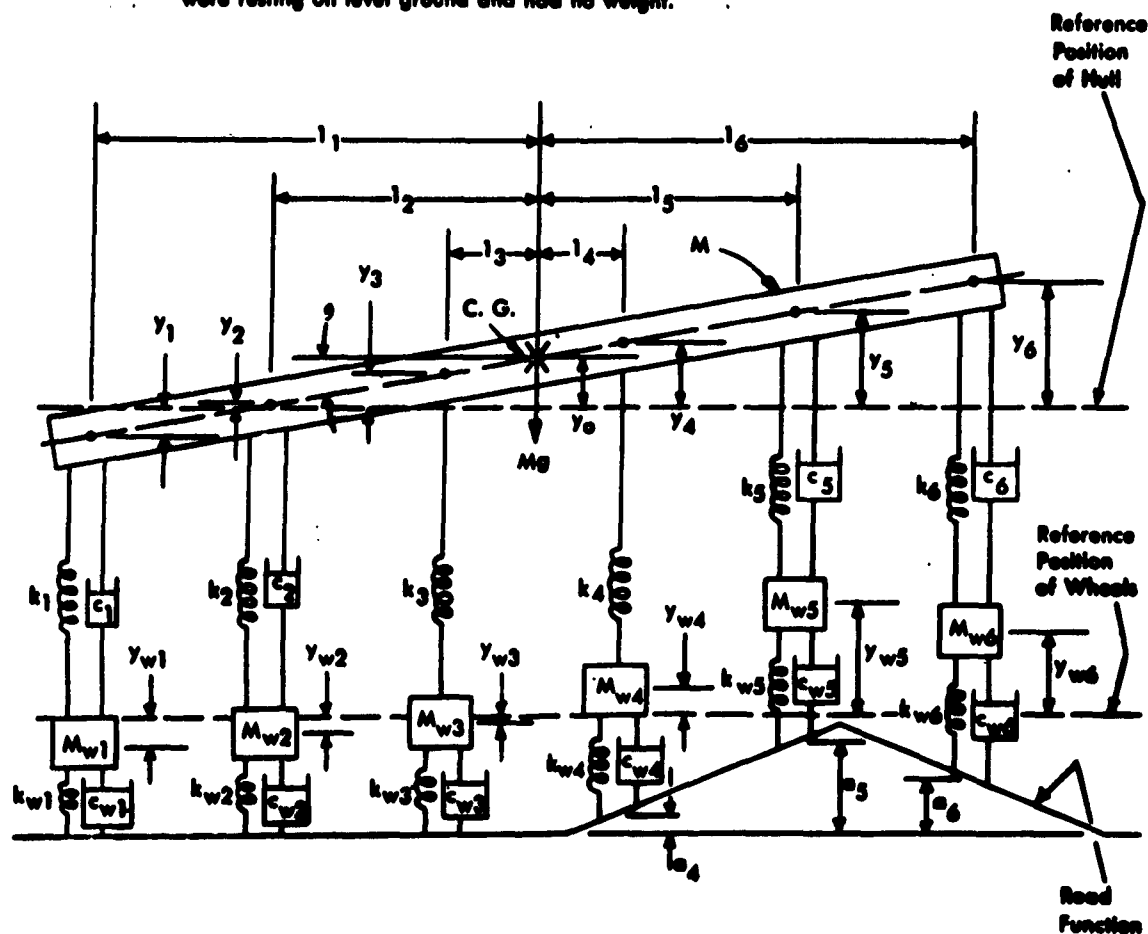


Fig. 7 SCHEMATIC REPRESENTATION OF THE TANK SUSPENSION SYSTEM

Data pertinent to the motion of the wheels and hull were recorded on motion pictures from which the pitch angle and vertical displacement of the hull and each of the road wheels could be measured.

Even a cursory analysis of the functioning of this suspension system makes it obvious that to take into account all effects would require a computer setup of formidable size with increased time and cost. To keep the simulation within reasonable limits, it was therefore desirable to look for simplifying assumptions which would not seriously affect the accuracy of the results. To accomplish this, the tank was considered as a solid body in space which has two degrees of freedom. The two degrees of freedom which were considered in the simulation of the M47 tank for the analog computer were:

1. Vertical motion of the center of gravity (usually called bounce).
2. Rotational motion (usually called pitch) about a transverse axis running horizontally through the center of gravity of the tank.

In the simulation (Figure 7), the hull of the tank was considered to be a rigid mass coupled to six masses (wheels) by six springs (torsion bars plus volute springs) and four dampers (shock absorbers). These six masses were, in turn, supported above the road by six springs and dampers representing the spring and damping effect of the rubber tires and track.

Only six wheels and half the mass of the tank were considered since it was assumed that the tank is symmetrical about a longitudinal vertical plane.

In general, it is convenient to start with an oversimplified simulation and make further refinements with their attendant complications as the need becomes evident. The major assumptions which were finally found to be satisfactory were not all determined in advance, but grew out of the correlation work itself. Careful observation of the test results gave clues to the factors which should not be neglected for realistic simulation. For example, it was convenient in the initial computer studies to assume that the wheels did not leave the ground. However, observations from the motion picture records clearly established the importance of simulating this effect.

The assumptions concluded and used are as follows:

- (1). The track has only minor effects on the hull response and if necessary may be neglected in the simulation. The only effect of the track included in the simulation was the modification which results from the action of the track as it affected the vertical motion of the road wheels when they moved onto and off the timber.
- (2). The spring effects of the torsion bars and the bump stops act in a vertical direction only and

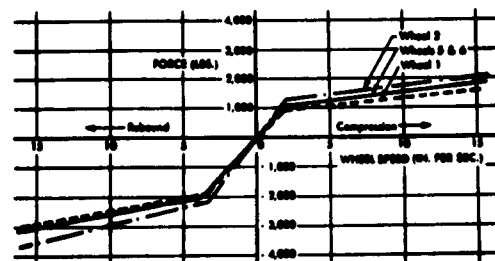
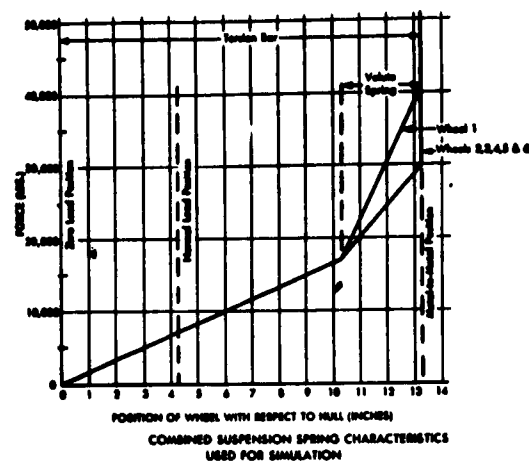


FIG. 8 SHOCK ABSORBER CHARACTERISTICS AS USED FOR SIMULATION

are represented by a curve of force versus deflection made up of a series of straight-line segments as shown in Figure 8.

(3). The shock absorber characteristics can also be represented by vertical components only, using curves made up of a series of straight-line segments as also shown in Figure 8.

(4). The horizontal forces on the tank can be neglected.

The analog computer determined the motion of the hull in pitch, and the vertical deflection, as functions of time by making use of equations expressing the following mathematical relations.

(1). The sum of the vertical forces on the hull is equal to the mass of the hull times its vertical acceleration. The vertical forces are those due to the torsion bars, the volute springs, the shock absorbers, and gravity.

(2). The sum of the moments about the pitch axis through the center of gravity is equal to the moment of inertia of the hull with respect to its center of gravity times the angular acceleration of the hull.

(3). The deflection of each road wheel with respect to the hull is a function of (a) the vertical position of the road wheel, (b) the vertical position of the

1. Vertical motion of the ball.

$$\begin{aligned}
 M\ddot{y}_0 = & -k_1(y_1 - y_{w1}) - k_2(y_2 - y_{w2}) - k_3(y_3 - y_{w3}) - k_4(y_4 - y_{w4}) \\
 & - k_5(y_5 - y_{w5}) - k_6(y_6 - y_{w6}) - c_1(\dot{y}_1 - \dot{y}_{w1}) - c_2(\dot{y}_2 - \dot{y}_{w2}) \\
 & - c_3(\dot{y}_3 - \dot{y}_{w3}) - c_4(\dot{y}_4 - \dot{y}_{w4}) - c_5(\dot{y}_5 - \dot{y}_{w5}) - c_6(\dot{y}_6 - \dot{y}_{w6}) + Mg
 \end{aligned}$$
2. Vertical motion of wheel 1.

$$\begin{aligned}
 M_{w1}\ddot{y}_{w1} = & k_1(y_1 - y_{w1}) - k_{w1}(y_{w1} - a_1) + c_1(\dot{y}_1 - \dot{y}_{w1}) \\
 & - c_{w1}(\dot{y}_{w1} - \dot{a}_1) + M_{w1}g
 \end{aligned}$$
3. Vertical motion of wheel 2.

$$\begin{aligned}
 M_{w2}\ddot{y}_{w2} = & k_2(y_2 - y_{w2}) - k_{w2}(y_{w2} - a_2) + c_2(\dot{y}_2 - \dot{y}_{w2}) \\
 & - c_{w2}(\dot{y}_{w2} - \dot{a}_2) + M_{w2}g
 \end{aligned}$$
4. Vertical motion of wheel 3.

$$M_{w3}\ddot{y}_{w3} = k_3(y_3 - y_{w3}) - k_{w3}(y_{w3} - a_3) - c_{w3}(\dot{y}_{w3} - \dot{a}_3) + M_{w3}g$$
5. Vertical motion of wheel 4.

$$M_{w4}\ddot{y}_{w4} = k_4(y_4 - y_{w4}) - k_{w4}(y_{w4} - a_4) - c_{w4}(\dot{y}_{w4} - \dot{a}_4) + M_{w4}g$$
6. Vertical motion of wheel 5.

$$\begin{aligned}
 M_{w5}\ddot{y}_{w5} = & k_5(y_5 - y_{w5}) - k_{w5}(y_{w5} - a_5) + c_5(\dot{y}_5 - \dot{y}_{w5}) \\
 & - c_{w5}(\dot{y}_{w5} - \dot{a}_5) + M_{w5}g
 \end{aligned}$$
7. Vertical motion of wheel 6.

$$\begin{aligned}
 M_{w6}\ddot{y}_{w6} = & k_6(y_6 - y_{w6}) - k_{w6}(y_{w6} - a_6) + c_6(\dot{y}_6 - \dot{y}_{w6}) \\
 & - c_{w6}(\dot{y}_{w6} - \dot{a}_6) + M_{w6}g
 \end{aligned}$$
8. Pitch of the ball.

$$\begin{aligned}
 J\ddot{\theta} = & -l_1k_1(y_1 - y_{w1}) - l_2k_2(y_2 - y_{w2}) - l_3k_3(y_3 - y_{w3}) \\
 & + l_4k_4(y_4 - y_{w4}) + l_5k_5(y_5 - y_{w5}) + l_6k_6(y_6 - y_{w6}) \\
 & - l_1c_1(\dot{y}_1 - \dot{y}_{w1}) - l_2c_2(\dot{y}_2 - \dot{y}_{w2}) + l_5c_5(\dot{y}_5 - \dot{y}_{w5}) \\
 & + l_6c_6(\dot{y}_6 - \dot{y}_{w6})
 \end{aligned}$$
9. $J = Mh^2$
10. $y_1 = y_0 + l_1\theta$
11. $y_2 = y_0 + l_2\theta$
12. $y_3 = y_0 + l_3\theta$
13. $y_4 = y_0 - l_4\theta$
14. $y_5 = y_0 - l_5\theta$
15. $y_6 = y_0 - l_6\theta$
16. $\dot{y}_1 = \dot{y}_0 + l_1\dot{\theta}$
17. $\dot{y}_2 = \dot{y}_0 + l_2\dot{\theta}$
18. $\dot{y}_3 = \dot{y}_0 + l_3\dot{\theta}$
19. $\dot{y}_4 = \dot{y}_0 - l_4\dot{\theta}$
20. $\dot{y}_5 = \dot{y}_0 - l_5\dot{\theta}$
21. $\dot{y}_6 = \dot{y}_0 - l_6\dot{\theta}$

FIGURE 9 - EQUATIONS USED IN THE SIMULATION

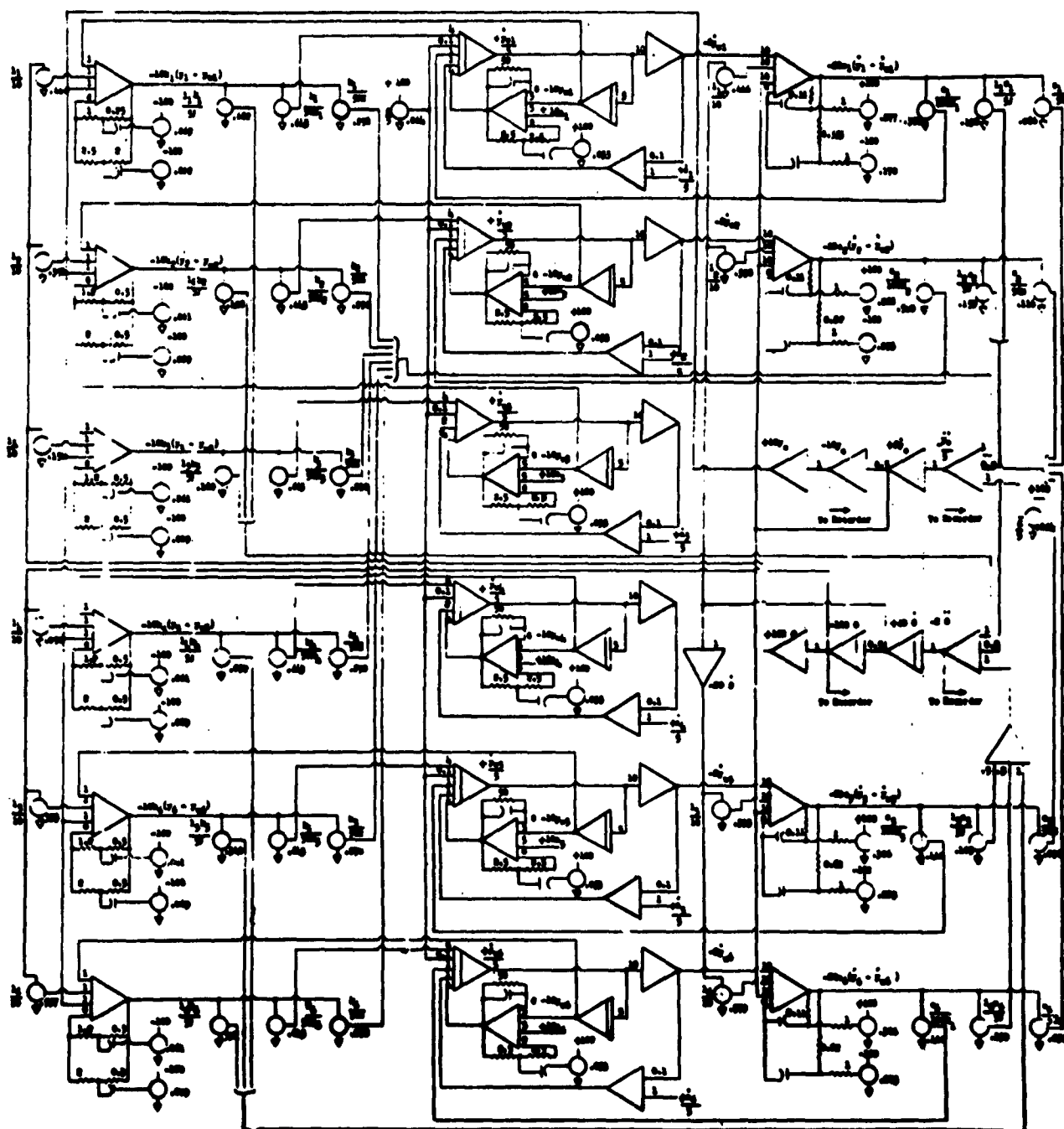


Fig. 28 SCHEMATIC DIAGRAM OF SUSPENSION SYSTEM SIMULATION

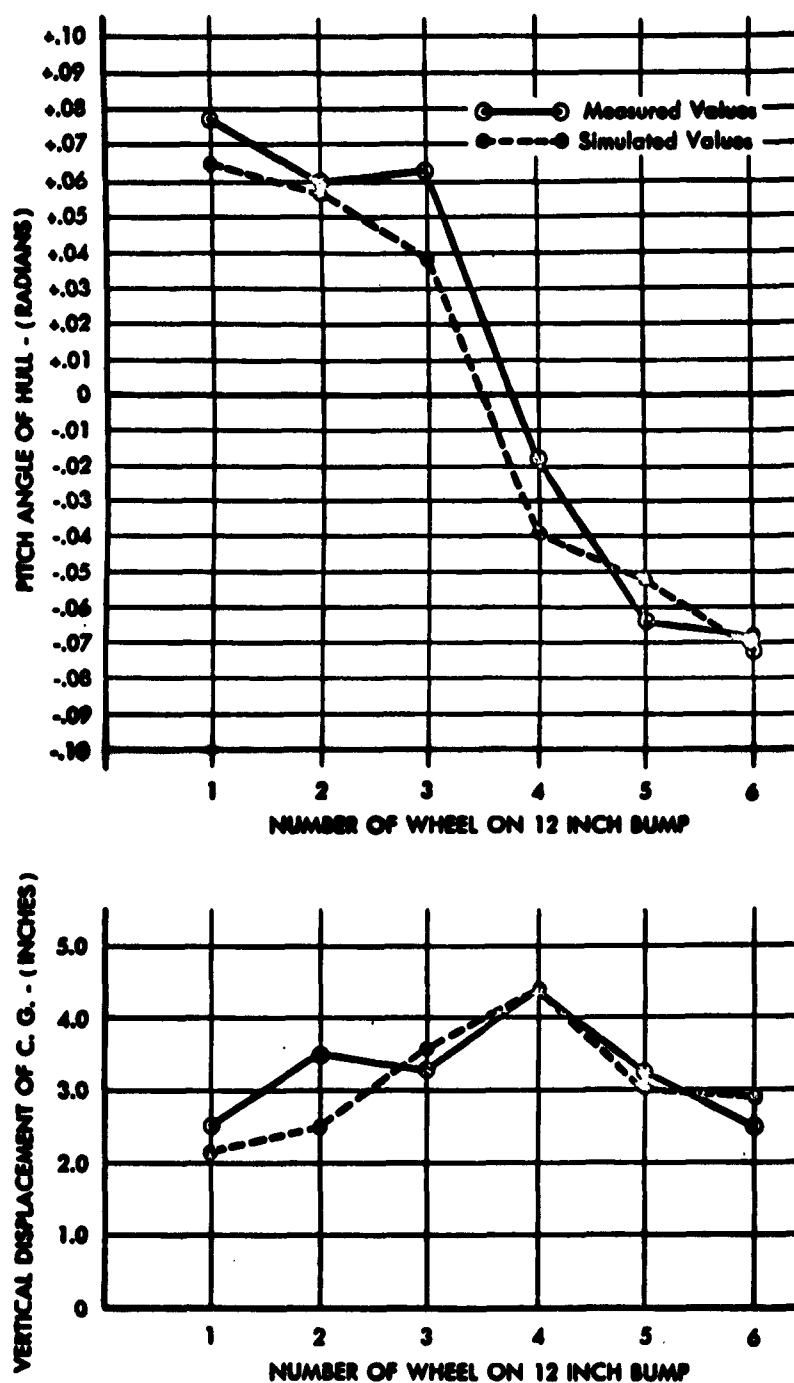


Fig. 11 COMPARISON OF MEASURED AND SIMULATED VALUES OF PITCH AND VERTICAL DISPLACEMENT OF HULL UNDER STATIC CONDITIONS

center of gravity of the hull, (c) the pitch angle of the hull, and (d) the known geometry of the suspension.

(4). The summation of vertical forces on each road wheel is equal to the mass of the wheel times its vertical acceleration. The vertical forces on each wheel are those due to the torsion bar, the volute spring, the shock absorber and the reaction due to the spring and the damping properties of the rubber in the track and tire.

The equations resulting from the above mathematical relations are shown in figure 9, and figure 10 shows the schematic diagram of the vehicle simulation.

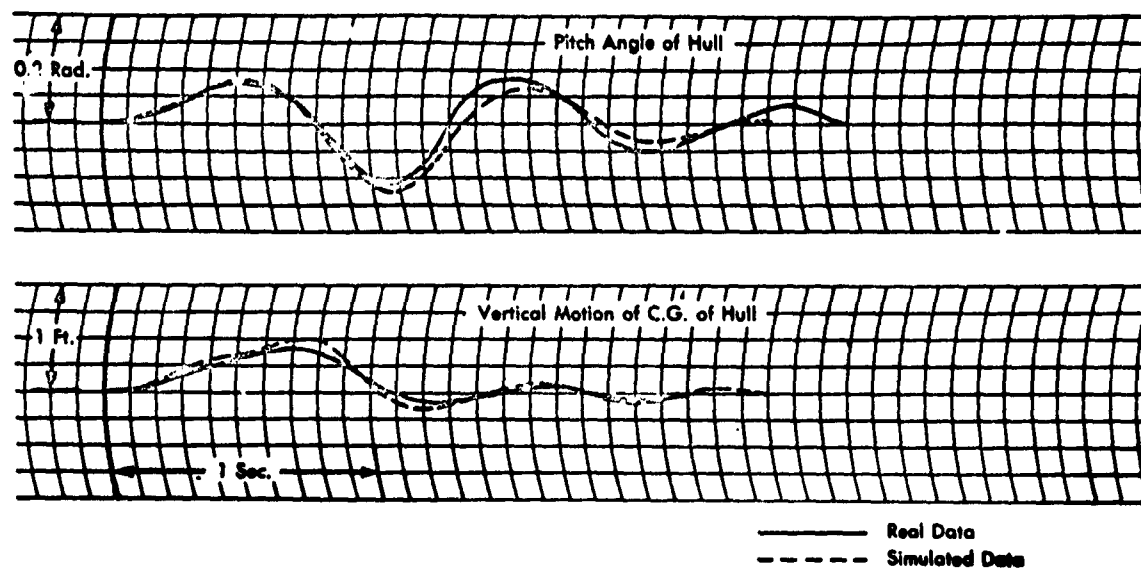
As a means of comparing the real tank with the simulated tank, under the least complicated conditions, checks were made of hull pitch and vertical deflection under static conditions. These data were obtained by riding the tank unto an obstacle, allowing it to come to rest, and then measuring pertinent distances, from which vertical displacement and pitch of the hull could be determined. Comparing these measured quantities with the simulated values provides a check on the validity of the computing technique for the steady-state conditions. This comparison is shown for the case of the 12-inch square obstacle in Figure 11, where pitch and vertical displacement are plotted for different horizontal locations

of the tank with respect to the bump. It can be seen that the correlation between real and simulated values is, for the most part, within the limits of the measurement errors.

An example of the correlation of the dynamic response of the tank for runs over the inverted V ramp at 20 fps is shown in Figure 12. The correlation of the response of the tank for runs over the 12" square timber is also shown in Figure 12.

In addition to the work on tank suspension systems, considerable effort has been expended on the simulation of tank gun control systems and components. Special attention has been given to the problem of stabilizing the tank gun in space. It has been possible to predict with reasonable accuracy the performance of a given gun stabilization system when coupled to the aforementioned suspension simulations. The addition of a pictorial presentation of target and sight-field, along with the necessary controllers, permits the human gunner to be included in the overall simulation. A simulation such as this allows the vehicle to be evaluated as a complete, dynamic fighting unit. Such an evaluation is of much greater value than attempting to evaluate the individual components alone.

Another problem in which the analog computer has been useful is in the suspension of high-inertia guns and turrets on moving hulls. In cases where gun or turret inertia is



REAL vs. SIMULATED RESPONSE OF M47 (8 Inch
Inverted - V Ramp, Tank Speed of 20 fps.)

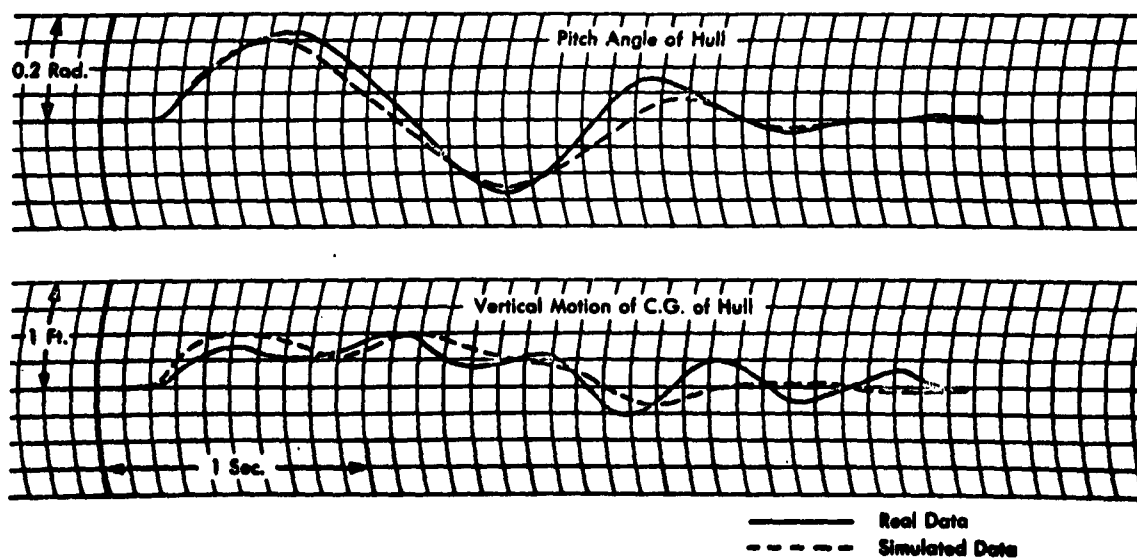


Fig. 12 REAL vs. SIMULATED RESPONSE OF M47 (12 Inch Square,
Tank Speed of 11 fps.)

extremely high, it becomes difficult to "hold" the gun or turret relative to the hull while moving over rough terrain. This problem has been overcome by the use of resilient members between hull and gun. The design of these resilient members is closely tied in with the vehicle suspension and gun control system. Simulation of the suspension, gun, and resilient mount has allowed the determination of design parameters which would otherwise have had to be determined from vehicular tests.

Experimental work is also in progress to provide the human gunner a movable seat arrangement which will accept inputs from the simulation and impart the vehicle motion to the gunner. It is recognized that the effectiveness of the gunner (and other crew members) is affected by vehicle motion. The ability to study these effects under controlled conditions will be of great value.

Also planned is a function generator to impart a realistic terrain function to the simulation. Topographical maps of various terrains in different parts of the world where tank warfare is contemplated would be reproduced and studied.

A word of caution may be appropriate. The analog computer is an instrument capable of solving numerous simultaneous equations which result from effectively isolating each component system, and usually, the equations for such component systems are of the second order and the number of these

equations is of no consequence because the computer will eliminate the time necessary to solve them. However, the application of an electronic computer to the solution of an engineering problem depends on the availability of numerical data. Hence, effectiveness of the computer approach will depend on the completeness and validity of these data. This is not as disadvantageous as it may seem. Let us assume that a radically new system is to be analyzed, that the information may be incomplete, and requires extensive extrapolation. In this case the computer will provide qualitative results as to the amount and manner in which each design variable affects the system's performance. The simulation will indicate the design of the system which has optimum characteristics, but will not provide an absolute measure of performance.

The foregoing research program verified that the results obtained by the computer were valid and within limits for the design and development of problems involved in various types of wheeled and tracked vehicles.

While it is true that computer techniques will save an enormous amount of time and money which normally would be spent in construction cost and laboratory tests, much more could be accomplished if more engineers were adept in correlating the physical reaction of structures in terms of graphical data. This difficulty was encountered and was solved by developing a "simuscope". This is a visual aid which presents a reproduction to scale of the motion of the hull and wheels as the vehicle goes over an obstacle at different speeds.

In view of the claims which have been made, one wonders why computers are not used more frequently. The reasons are:

1. A lack of knowledge of just what computers can do to solve problems in a particular engineering field. The automatic computing field is a relatively new one and the "word" has not been passed on to everyone.
2. A mistaken idea as to the cost of computing equipment. Engineers who have not had time to become familiar with computers feel that it takes "a million dollar machine" to solve problems. This is not the case at all. If one were to hazard a guess on the cost of a machine to solve a particular problem it would probably go something like this: where computers are applicable, the cost of the computing equipment would be much less than one costly "guestimate" where the entire design has to be scrapped after the pilot model is manufactured, and in many cases the cost is less than changing major components in the experimental systems.

It can be expected that as engineers become aware of the numerous applications of the computer in their respective fields, they will begin to use them to fuller advantage. Perhaps as the novelty of the equipment wears off and the word "computer" ceases to be synonymous with mystery and complexity the analog computer will command widespread use in engineering organizations of both government and industry.

PROBLEMS IN THE APPLICATION OF LEAST SQUARES

TO FUNCTION GENERATOR DESIGN

By John G. Tappert

A function generator is a circuit or network used in an analogue computer to store, and exhibit on demand, a predetermined relationship between a number of variables. Consider the relationship,

$$f(y, x_1, x_2, \dots, x_n) = 0.$$

A function generator used to solve this relationship would have the property of exhibiting, in analogue form, the value of y which satisfies the given equation whenever a set of x 's are introduced in their appropriate analogue form. Normally, it is required that the relationship be satisfied only in a limited domain and within a specified accuracy. Further, in computed applications, the functional relationships are limited to those which are single valued and have finite first derivatives of the dependent variable.

Function generators are essential to many military computers. Practical design of such circuits requires that efficient use be made of available computing elements. One of the most simple, yet extremely versatile computing elements is the function of a single variable. That is, an arbitrary function of a single variable is, in most cases, easily and reliably instrumented to a high degree of accuracy. It is reasonable, then, in designing a function generator involving many variables to attempt to approximate the desired functional relations by combinations of functions of single variables. This paper will deal with the mathematical problems involved in the determination of optimum functions of single variables in higher order curve fitting.

The techniques for applying the principle of least squares to curve fitting problems, in the case where a finite number of parameters are to be optimized, are well known. If the residuals can be expressed as linear functions of the

parameters to be determined, then these parameters can be determined by solving a set of linear equations. In our present problem, the unknowns are functional relationships rather than coefficients of known functions. However, the former is readily reduced to the latter by defining the functional relationship in terms of a finite number of parameters.

Let us consider an example along with a typical set of conditions. Let there be a relation

$$y_1 = f(x_1, x_2, x_3), \quad (1)$$

which is to be instrumented in the form

$$y_2 = F_1(x_1) + F_2(x_2) + F_3(x_3). \quad (2)$$

It is desired to determine the functions F_1, F_2, F_3 which minimize

$$\iiint (y_2 - y_1)^2 dx_1 dx_2 dx_3 \quad (3)$$

over some specified domain in X-space. However, no explicit analytic expression for (1) is known; but it is possible, at considerable cost in time and labor, to determine sets of data which satisfy the relation. Further, the procedure employed in deriving these sets of data does not permit direct choice of specific values of x's or y; thus, the available data is irregularly spaced in all variables.

It might seem reasonable, under these conditions, that if we took a representative set of data and replaced the integral (3) by a summation, the problem could be solved. But suppose, for example, that among the selected data points no two values of say, x_1 were identical. Then, as far as the data are concerned, y_1 could unambiguously be represented as a function of x_1 alone, and a solution with all the residuals vanishing would be obtained by making F_2 and F_3 constants. Such a solution obviously would, in general, not satisfy the condition (3) and therefore would not be satisfactory.

These conditions, then, are not sufficient to solve the problem. It is necessary, as it always is when we expect to define continuous functions by a finite number of points, to assume that the functions have that rather ill-defined characteristic of being "well behaved". For our present purpose, we can say that function is "well behaved" if it has no structural peculiarities that are not revealed by the available data. This property of "well-behavedness" is akin to the property of continuity in that it is a local characteristic, involving only the relation of data points in a finite but limited region.

We are, then, faced with a two-fold problem:

- a. To define a functional relationship in terms of a finite number of parameters.
- b. To define a functional relationship in such a manner that it is locally "well-behaved" at every point, without imposing restrictions on its gross shape.

There are, undoubtedly, many ways in which these two conditions can be met reasonably well. The method that has been employed successfully at Frankford Arsenal consists of defining the function in terms of a finite number of cardinal values and a low order interpolation law. The procedure in defining a function of x would be as follows:

Normalize the variable, x , so that its excursion is bounded by convenient cardinal numbers, say, 0 to 10, and define the value of the function at all the cardinal points of the argument by a parameter. Then the function would be completely defined by

$$F(N + k) = \sum a_n F_n.$$

where the F 's are the cardinal value parameters, the a 's are the interpolation coefficients, and the argument is written as a cardinal number, N , plus a decimal fraction, k . In the case of Lagrangian interpolation, the numerical values for the a 's depend on k only and their assignment as coefficients of

specific F 's depends only on N .

This procedure obviously defines a functional relationship in terms of a finite number of parameters; the low order interpolation law adequately controls the "well-behavedness" of the function; and the fact that with low order interpolation all the coefficients vanish except those for the cardinal values in the immediate neighborhood of the argument limits the restrictions imposed by this definition on the gross shape of the function.

The problem of optimizing functional relationships can, therefore, be reduced to the problem of optimizing constants, and this problem in turn is amenable to solution by known conventional methods.

Experience has shown that difficulties can arise if certain precautions are not taken. Some of these are discussed below.

It is very desirable to define the mathematical problem in such a manner that the solution is unique. Consider a structure

$$y = f_1(x_1) + f_2(x_2),$$

and let f_{10} and f_{20} be functional relations which result in a sum of the squares of the residuals which no other functions can reduce. Then there are a whole set of functions,

$$\begin{aligned} f_1(X_1) &= k f_{10}(X_1), \\ f_2(X_2) &= \frac{1}{k} f_{20}(X_2), \end{aligned}$$

which give identically the same residuals. These solutions are, of course, identical as far as a computer designer is concerned, but are different solutions mathematically, and therefore leave things indeterminate. This type of difficulty can usually be eliminated by reducing the number of parameters by one, and defining the function at one of the cardinal values in terms of another,

$$f_1(N_2) = f_1(N_1) + \text{constant}.$$

This fixes the scale of the f_1 function without restricting its shape, and eliminates the multiplicity of equivalent solutions. In more complicated structures, the possibilities of such multiple solutions may be less obvious, and their elimination, without imposing unnecessary restrictions, can call for considerable ingenuity.

Every effort should be made to reduce to a minimum the number of undetermined parameters required to define a function generator structure. This not only reduces the amount of labor required, but also provides the best insurance that the solution is not unduly influenced by the particular finite selection of data which are taken to represent a bounded continuum.

A designer, in setting up a structure to be investigated, usually has a fair idea of the general form of the functions which will be required for it. Say, for example, that there is reason to believe that one of the functions will turn out to be something like the tangent of an angle, between 0° and 85° . That is,

$$f(x) \approx \tan x.$$

To adequately represent such a function with a low order interpolation law would require closely spaced parameter values in the region of large values of x , but could tolerate wide spacing in the region of small values of x . Considerable economy of parameters can be realized by a change of variable,

$$f(x) \approx g(\tan x).$$

Since the g function would be almost linear, it presumably could be adequately defined by a small number of parameters and still retain the advantage of a low order interpolation law. Alternative ways of handling this, when the change in variable is not convenient, amount to breaking up the undetermined function into two terms or factors, one of which is predetermined. Thus, there is no loss in generality in defining

$$f(x) = \tan x + g_1(x),$$

or

$$f(x) = \tan x \cdot g_2(x).$$

Here, again, fewer parameters would be required to adequately define g_1 or g_2 than would be required to define f .

SOME TYPICAL PROBLEMS CONFRONTING MATHEMATICIANS AT BRL

By Saul Gorn

INTRODUCTION

At two other meetings sponsored by the Office of Ordnance Research, namely the Symposia of April 1953 and April 1954 in New York and Chicago respectively, there were talks on subjects similar to the one I am called upon to discuss today. Since Dr. Lotkin, who gave the address in New York, and Professor A. A. Bennett, who spoke in Chicago, were both, like myself, connected with the Computing Laboratory of the Ballistic Research Laboratories at Aberdeen Proving Ground, it might seem that they have suddenly discovered some big gap in their discussions for me to expand upon.

This they definitely did not do.

And yet I am sure that they would both agree that the automatic, high speed, digital computers have had such a dramatic effect on both our ability to solve our problems and our attitude on how to do so that this effect will bear describing again and again.

However, you will understand why, in this description, I choose different examples from theirs.

Both referred to the variety of machines, including Eniac, Edvac, Ordvac, etc., and Dr. Lotkin described their main characteristics. Professor Bennett described the variety of application and Dr. Lotkin gave their mathematical categories. Finally, Professor Bennett gave the percentage of computing time devoted to each class of application and spoke about the effect of high speed computers on mathematics itself.

I want to touch upon the sources and nature of the requests for our services.

Within BRL we get problems from the Exterior Ballistics Laboratory, The Interior Ballistics Laboratory, The Ballistic Measurements Laboratory, The Terminal Ballistics Laboratory, The Weapon Systems Laboratory, and problems from the Computing Laboratory itself.

Almost one third of the projects are concerned with computations of firing and bombing tables. From exterior ballistics come many data reduction problems involving smoothing, least square fitting, and the like, as well as a number of air flow problems. Interior ballistics has been yielding problems in the determination of trajectories, burning rates, and molecular structures. Beside numerous data reduction problems, ballistic measurements give rise to various theoretical problems concerned with measuring systems. We will see an example in a moment. Terminal ballistic problems give rise to a host of theoretical studies involving elasticity, penetrability, fragmentation, and distribution of fragments.

The problems of weapon systems evaluation are largely concerned with probability, statistics, and game theory. Finally, in addition to the firing and bombing tables, The Computing Laboratory has many problems, mathematical and logical, not directly tied to specific computations, concerned with efficient use of the high speed computers. More will be said on this subject later.

The sources of problems outside BRL are the arsenals, other Government installations, Government contractors, and universities. From the arsenals we have been getting large data reduction problems, often by complicated least squares procedures, and computations of complicated trajectories; other problems from Government installations and Government contractors have involved computations of complicated multiple integrals, high precision solutions of complicated numerical equations, boundary value problems, and solutions of medium sized systems of linear equations. Problems from pure mathematical research come to us by way of research contracts awarded by the Office of Ordnance Research itself. An example of this type will also be touched upon shortly.

AN INTERIOR BALLISTICS PROBLEM

An example of how an advanced mathematical viewpoint can aid in the choice of an economical computing method by examining the variables in a range in which they are physically meaningless is the following. The interior ballisticians, in studying the motion of a projectile within a gun after the charge has burned out, have a physical model leading to the trajectory equation:

$$u u' = x^{\sqrt{}} - u^{-\frac{1}{2}}$$

where $\sqrt{}$ is a constant in the neighborhood of 1.3 and x and u are dimensionless quantities related respectively to the free volume behind the projectile and the projectile velocity. It is desirable to have a table of solution curves for various sets of initial conditions corresponding to different guns, different charges, different projectile weights, etc. From the very statement of the problem it is evident that the equation must be physically meaningless in a neighborhood of $x = 0$, and, indeed, the computed values are only of physical interest in the range

$$\begin{aligned} .05 \leq x \leq 4 & \quad \sqrt{ } = 1.25, 1.3 \\ .2 \leq u \leq 4 & \end{aligned}$$

Nevertheless, to choose an efficient organization of the computation calls for a view "in the large" of all the solutions in the x, u - plane, whose distribution is greatly influenced by the nature of the singularity at the origin. By an overall analysis of the type much used in non-linear mechanics one can quickly sketch the family of solutions from the nature of the singularity, the zero isocline, the infinite isocline, and the locus of inflections; all these may be determined directly from the equation. The situation is roughly summarized in the accompanying diagram:

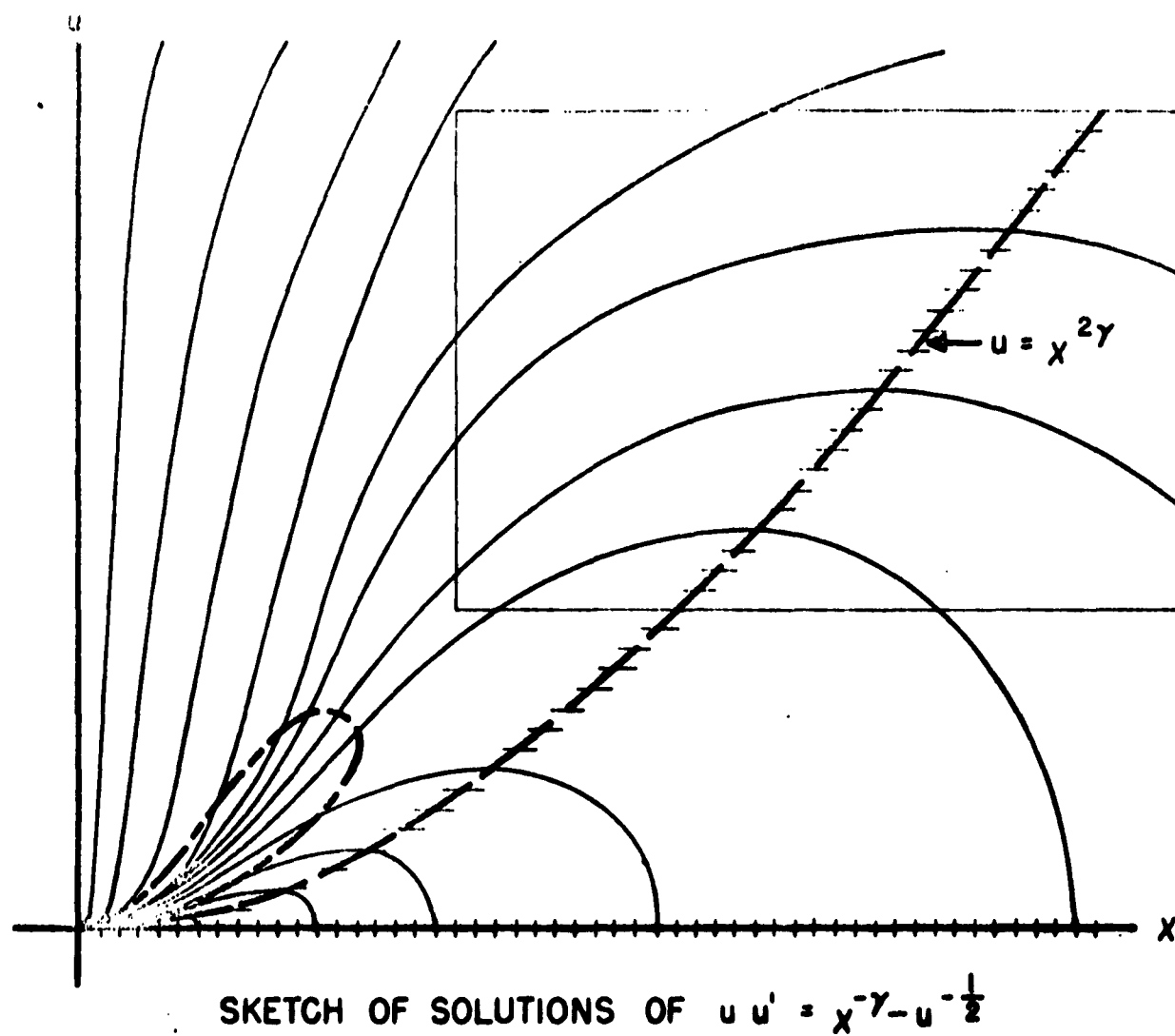


FIG. 1

Because of this view of the solutions, it was decided to organize the numerical solution by machine by 1. - beginning each solution at $u = x^2 y$, proceeding to the right until $u' = -1$. 2. - Switching variables and continuing on to the edge of the desired region, 3. - proceeding left from the same initial point in the same way switching when $u' = 1$. This procedure was carried out automatically.

A BALLISTIC MEASUREMENTS PROBLEM

The physicists concerned with ballistic measurements sometimes measure a distance by sending out a radio signal, frequency-modulated by a number of different tones. The signal is automatically returned from the object whose distance is being measured, and the phase differences in the different modulating frequencies, also called range channels, are used like the dial readings on a gas meter to give, so to speak, successive decimal digits in the distance. When, however, there are superfluous return signals, due to such things as unwanted reflections, a distortion of these range channels is the result, causing an error in the estimate of distance. To get an idea of the magnitude of such errors, a simplified model with one unwanted reflection was assumed, and the reflected signal was supposed to have a strength as large as that of the direct signal itself.

The Computing Laboratory had, then, to produce accurate answers to the following:

Find the phase shifts, α_1 and α_2 , for two values of the deviation ratio, β , and eight values each of θ_0 , v , and ξ , having the following meanings -

- θ_0 = carrier phase lag at the receiver of the reflected wave relative to the direct wave,

v = a quantity proportional to the path difference

ξ = a quantity connected with phase lag between modulating frequencies,

where α_1 and α_2 are given by:

$$\tan \alpha_1 = G_1/H_1, \quad \tan \alpha_2 = G_2/H_2,$$

$$G_1 = \frac{1}{2} \sin 2\nu + \frac{1}{\beta} \sum_{i=1}^F (m_{i-1} - m_i) \sin (B_i + \nu),$$

$$H_1 = \frac{1 - \cos 2\nu}{2} + \frac{1}{\beta} \sum_{i=1}^F (m_{i-1} - m_i) \cos (B_i + \nu),$$

$$G_2 = \frac{1}{2} \sin 20\nu + \frac{1}{10\beta} \sum_{i=1}^F (m_{i-1} - m_i) \sin (10B_i - \xi + 10\nu),$$

$$H_2 = \frac{1 - \cos 20\nu}{2} + \frac{1}{10\beta} \sum_{i=1}^F (m_{i-1} - m_i) \cos (10B_i - \xi + 10\nu),$$

where $-\pi < B_i < \pi$, and

$B_i, m_i, i = 1, \dots, F$ are such as to yield

$$\psi = (2m+1)\pi \quad \text{for}$$

$$\psi = 0 - 2\beta [\sin \nu \cos B + \sin 10\nu \cos (10B - \xi)].$$

There are essentially numerical equations, although the intermediate variables M_i must be integers. For different assigned integers for M , the last equation must be solved for B to yield a corresponding B_i . The main problem is essentially one of organizing the arrangement of solution of somewhere in the order of 10,000 such equations in such a way that the answers are properly grouped to find G_1 , H_1 , G_2 , and H_2 and hence α_1 and α_2 . This combination of bookkeeping and traffic control with computed intermediate results is one of the main capabilities of the automatic high speed machines. The numerical methods adopted would have taken a skilled and tireless hand computer at least one man year. Less than thirty hours of machine time was actually involved.

A STATISTICAL PROBLEM

In evaluating the effectiveness of a gun and ammunition it is customary to use the standard deviation from the mean point of impact, both in range and deflection. Sample variances or standard deviations are customarily used to estimate the corresponding population parameters. However, in order to reduce the effect during range firing of shifts in the mean point of impact due to wind, A. A. Bennett around 1919 proposed that successive differences in the observed ranges and deflections be used to obtain estimates of the aforementioned parameters. This has led to the problem of computing the probability distribution of such sample statistics as:

$$\sigma_{n,p}^2 = \frac{d_{n,p}^2}{S_n^2}, \text{ where}$$

$$d_{n,p}^2 = \frac{1}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^n (\Delta^p X_i)^2, \quad (2 \leq p \leq n-1), \text{ and}$$

$$S_n^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2, \quad (n = 2, 3, \dots).$$

Here, of course,

$$\Delta^p X_i = \Delta^{p-1} X_i - \Delta^{p-1} X_{i-1} = \sum_{j=0}^p (-1)^j \binom{p}{j} X_{i-j},$$

$$\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j,$$

and the X_i are independent random samples from a normal population.

It is desirable to have tables of these distributions for sample sizes, n , from 3 to 25, and order of difference, p , from 1 to 5. The distributions themselves were completely characterized by von Neumann as long ago as 1941; the problem is to choose a reasonable method of computing them.

At least nine methods, ranging from the naive-but-not-to-be-ignored to the sophisticated, have been considered. The method being investigated at present is somewhat typical in that it probably would not have been considered if the high-speed, automatic machines had not been available.

First an orthogonal transformation is applied to reduce the problem to that of computing the distribution of

$$\beta = \frac{B_1 y_1^2 + \dots + B_m y_m^2}{y_1^2 + \dots + y_{n-p}^2}, \quad n = n - p.$$

The characteristic values, B_1 , for the various n and p have been computed. Here

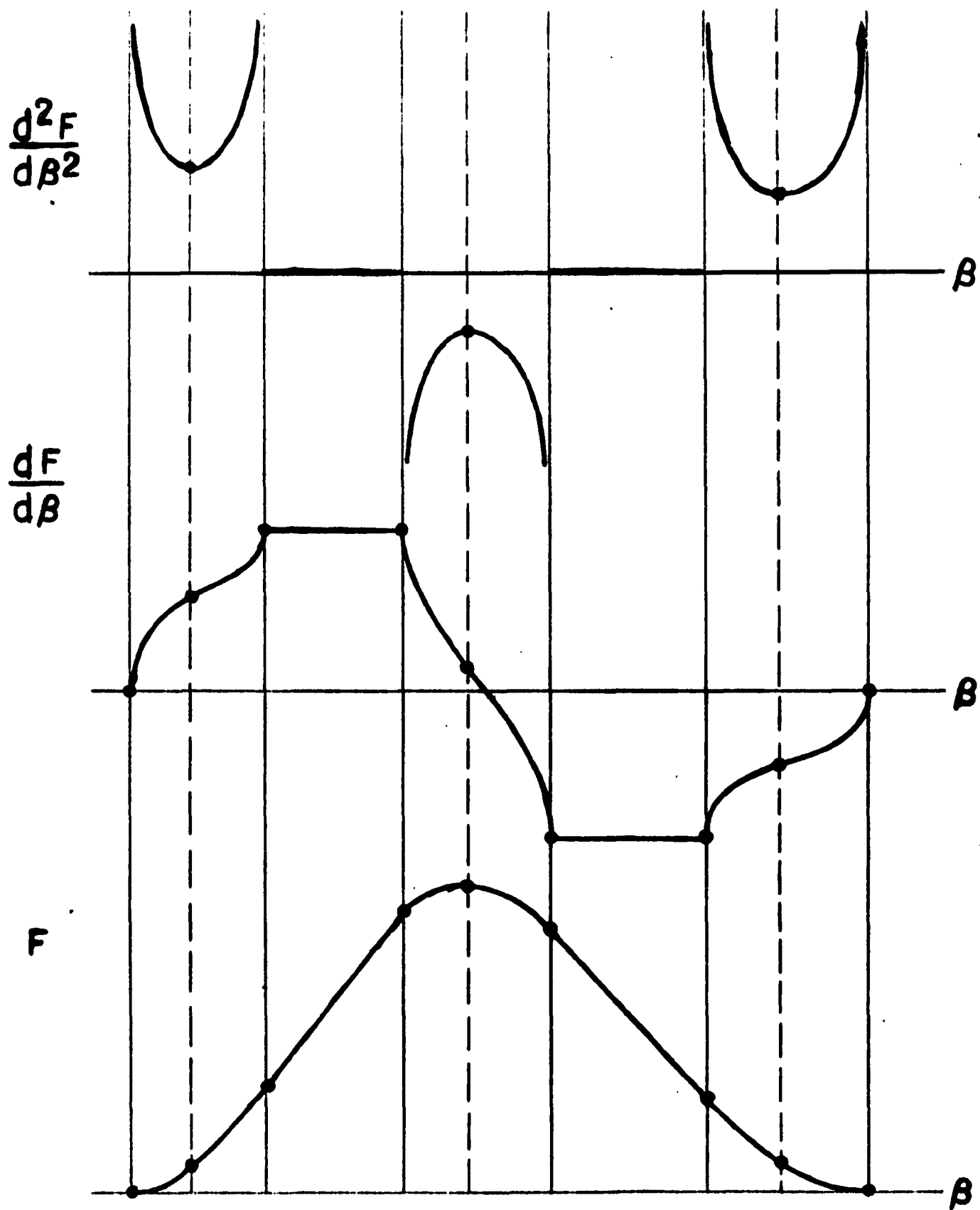
$$B_1 > B_2 > \dots > B_m.$$

The method under consideration begins with von Neumann's relation for the probability density, $F(\beta)$, applicable when n is even:

$$\frac{1}{(\frac{n}{2} - 1)!} \frac{d^{\frac{1}{2}n-1}}{d\beta^{\frac{1}{2}n-1}} F(\beta) = \begin{cases} 0 & \text{if } v \text{ is even} \\ \frac{(-1)^{\frac{n-v-1}{2}}}{\pi} \frac{1}{\sqrt{\prod_{v=1}^n (B_v - \beta)}} & \text{if } v \text{ is odd} \end{cases}$$

for $B_{v+1} < \beta < B_v$.

Thus a sketch of $F(\beta)$ and its $1/2n-1$ derivatives, when $n=6$, would look as follows:



SKETCH OF $F(\beta)$ AND DERIVATIVES FOR $m = 6$

FIG. 2

If one expands the non-zero expression for the $\frac{m}{\beta}$ - 1'st derivative of $F(\beta)$ into fractional power series about each of the singular points, B_v , the radius of convergence will be the smaller of $B_v - B_{v-1}$ and $B_{v+1} - B_v$.

Integrating these series term by term should permit us to advance along the β -axis by a method of analytic extension applied to all the derivatives simultaneously down to F . In all, this will call for the use of almost 7500 power series and almost as many polynomials, which should be handled automatically in interrelated groups to economize on the requirements for simultaneous storage of coefficients.

A PROBLEM FROM MATHEMATICAL RESEARCH

The following problem concerned with the theory of algebraic numbers came to us from the Office of Ordnance Research in support of a contract it had awarded. A number ξ , is called a cubic algebraic number if it satisfies an equation of the form:

$$x^3 - tx^2 + \bar{t}x - N = 0,$$

solutions, the conjugates of ξ , ξ_2 and ξ_3 ,

$$t = \xi_1 + \xi_2 + \xi_3 \text{ called the } \underline{\text{trace}} \text{ of } \xi,$$

$$N = \xi_1 \xi_2 \xi_3 \text{ called the } \underline{\text{norm}} \text{ of } \xi,$$

t , \bar{t} , and N rational numbers, and

ξ satisfies no such equation of lower degree.

If t , \bar{t} , and N are also integers, ξ is an 'integral algebraic number'. If, in addition $N = \pm 1$, ξ is called a unit. Like the numbers ± 1 , units divide exactly the other algebraic integers of their domain.

Professor H. Cohn of Wayne University designed an algorithm which within a cubic field produces many algebraic integers with small norms, and must yield at least one unit. It begins with a suitably chosen three by four matrix of algebraic numbers

$$\begin{pmatrix} P_1 & Q_1 & R_1 & S_1 \\ P_1' & Q_1' & R_1' & S_1' \\ P_1'' & Q_1'' & R_1'' & S_1'' \end{pmatrix}$$

with sign pattern

$$\begin{pmatrix} + & - & - & + \\ - & + & - & + \\ - & - & + & + \end{pmatrix}$$

and $P + Q + R + S = 0$

One can choose a column, add it to a second, and subtract it from a third. Of the twenty four possible results, between three and six will maintain the sign pattern, and the algorithm consists in branching out from these by the same process. However a matrix is the end of its line if some suitable rearrangement of its rows and columns makes its rows individually proportional to those of a predecessor. The factors of proportionality are units.

In order to make more modest demands on storage the sponsor replaced this branching algorithm by one whose results appeared in a linear order. The new algorithm was to be applied to the so-called 'cyclic cubic fields with prime conductor', beginning with matrices simply related to the solutions of:

$$w^3 + w^2 - mw - S = 0,$$

$$l = 3m + 1 \quad \text{a prime integer,}$$

$$S = \frac{(A+3)l-1}{27}, \quad \text{where}$$

$$A = 3k + 1 \quad \text{and} \quad g > 0 \quad \text{are}$$

$$\text{The unique solutions of } A^2 + 27g^2 = 4l.$$

It was found possible, using about a thousand storage positions on one of our machines to carry out all operations exactly except for those required to verify the sign patterns of the matrices. If the machines had more internal storage available, these sign discriminations could also have been made independent of approximations. As it was, automatic error controls were carried along with the computations.

The fact that the essential computations could be carried out to what is effectively infinite precision to produce some 3500 algebraic integers together with their determining equations for all primes up to 500 was made possible by the use of the so-called 'structure constants':

$$\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2 \quad \text{such that}$$

$$w_1^2 = \alpha_1 w_1 + \beta_1 w_2 + \gamma_1 w_3,$$

$$w_1 w_3 = \alpha_2 w_1 + \beta_2 w_2 + \gamma_2 w_3.$$

namely:

$$\alpha_2 = \frac{m+k+1}{3}, \beta_2 = \frac{g+m-\alpha_2}{2}, \gamma_2 = \frac{-g+m-\alpha_2}{2}$$

$$\alpha_1 = \alpha_2 - m - 1, \beta_1 = \frac{g-m-\alpha_2}{2}, \gamma_1 = \frac{-g-m-\alpha_2}{2}.$$

By use of these, every integer formed and its defining equation could be given as follows:

$$\frac{Y}{Z} = (a, b, c) = a w_1 + b w_2 + c w_3$$

a, b, and c rational integers,

$$t = -(a+b+c),$$

$$\bar{t} = -m t^2 + l(ab+bc+ca),$$

$$N = -S t^3 - l \{ \sigma_1 \sum a b^2 + \sigma_2 \sum a^2 b + abc \},$$

where

$$\sigma_1 = \frac{k-g+1}{2}, \sigma_2 = \frac{k+g+1}{2}$$

This problem thus had the interesting and paradoxical feature of including both infinite precision computations and automatic error analyses.

THE GENERAL PROBLEM OF ORGANIZING MACHINE USE EFFICIENTLY

The problems described so far, although diverse in source and background, have in common a kind of dynamic logic we have already referred to as bookkeeping and traffic control. They have all included the mechanization of decisions on how to continue, where the criteria are, of course, fixed. In this sense the machines can think, and the thinking can be on a large scale and very intricate, but must be routine as opposed to creative or executive. And in most problems, even though they are computational, a large part of the work is this making of routine decisions. This is even true when the problem is not, strictly speaking, computational, as, for instance, when we set the machines up to play war games. It is perhaps for this reason that programs for the machines are commonly called routines. One of the most surprising facts to emerge from constant attempts to extend the range of applicability of these machines is that vast portions of our judgments are routine and can therefore be mechanized. Perhaps these machines should not be called computers at all; they can do general information transforming of which there are other forms than the numerical one of computing; translating languages is an example.

An example of a large scale information transforming job already mentioned as a primary function of BRL's Computing Laboratory is the following: Given large quantities of range data, produce a firing table. The whole procedure can be thought of as one huge routine whose subroutines at present use different machines and groups of people. The people are used less and less within this routine and more and more in planning its mechanization. For example, a big part of the hand typing of the tables has been eliminated because the extraction of

the tabulated information is performed mechanically upon the computed trajectories and prepared mechanically for automatic typing. When machines with large internal storage become available to us, very little of the table production procedure will be left unmechanized; even checking for bad data can be made routine.

At present, then, more time is spent preparing problems for computation than it takes to do the computation. The problems are literally quicker done than said. Months are needed to prepare, against hours to compute. This is because much judgment is required in the preparation of routines.

And yet, as usual, a large portion of this judgment is itself routine, and can again be mechanized. We should be able to enlist the machines to help in preparing their own problems for their own solution, and thus possibly reduce this preparation time from months to minutes.

Let us pause, then, to break down this preparation time. When a problem is received, it must go through four stages before all the actual computations are performed. These may be called analysis, error analysis, coding, and code checking.

The analysis stage includes choosing the methods to be used, i.e. choosing the main subroutines in the routine together with their logical ordering, and doing such analytic things as say, substituting one power series into another. An example of the first type of analysis is the choosing of a difference equation approximation to a differential equation (the only strictly numerical operations these machines perform are addition, subtraction, multiplication, and division). This usually calls for executive judgment and would rarely be mechanized.

The second stage, error analysis, also has a large element of routine thinking in it. It should therefore be possible in many cases to arrange for the error analysis and control to be carried on along with the computation by the machine. We have mentioned this sort of thing in connection with the problem from mathematical research.

The coding stage, as everyone is aware, also contains a large element of the routine clerical type of thought, especially if well known methods are being applied. The routine element here consists in putting together the instructions themselves. But our machines store their orders exactly as they store their numbers, so that routine operations on their own orders can be carried out mechanically. This makes it possible, again if the storage is sufficiently extended, to have the machines automatically assemble their own routines from previously coded subroutines. Combining this possibility with the machines' capability of doing routine translation, this means that it is possible in a multi-machine installation such as BRL's Computing Laboratory to code problems in a universal abbreviated code, independent of the particular machine. The machines themselves would assemble their codes from previously coded

subroutines, and then translate the routines into their own code, making their own routine choices of storage assignments. This is a universal, semi-automatic coding. Experiments are being run on the spare time left by production commitments to see how expensive such a procedure would be in both storage and time.

The last stage, code-checking, can take considerable time since human errors in coding are almost bound to appear in long problems, such as some we have described. The existence of a large library of standard subroutines would do much to alleviate this situation; the existence of universal, semi-automatic coding procedure would do even more.

With these considerations in mind, it is, I believe, apparent why so much has been said about the effect these machines may have on the development of mathematics and on the type of personnel needed to program them.

REFERENCES

- M. Lotkin - BRL Memorandum Report No. 693 "Some Problems Solvable on Computing Machines" (Symposium in Applied Mathematics - 23-24 April 53)
- A. A. Bennett - "Some Numerical Computations in Ordnance Problems" (Symposium in Applied Math; 28, 29, 30 April 54)
- J. von Neumann - "Distribution of the Ratio of Mean Successive Difference to the Variance" - Annals of Math. Stat., 1941
- H. Cohn - "A Periodic Algorithm for Cubic Forms" - Am. Journ. Math. - Oct. 52
- S. Gorn and M. L. Juncosa - BRL Report No. 889 - "On the Computational Procedures for Firing and Bombing Tables"
- S. Gorn and R. Moore - BRL Report No. 893 - "Automatic Error Control, The Initial Value Problem in Ordinary Differential Equations"
- S. Gorn - BRL Report No. 816 - "On the Study of Computational Errors"

COMPUTING METHODS FOR THE SOLUTION OF
SYSTEMS OF DIFFERENTIAL EQUATIONS

By F. J. Murray

The Office of Ordnance Research has kindly given me an opportunity to discuss the work which I have done in connection with computing procedures for differential equations, and I am grateful for this invitation.

The experience of any individual is necessarily limited, and the methods which have interested me have been developed in connection with a certain range of problems and certain types of computing facilities. I must apologize then if this range is not coincident with your range of interest.

Most of this work arose as a consequence of consultation with the United States Navy Special Devices Center, and certain theoretical developments were carried on later under the auspices of the Office of Naval Research. Special Devices has had a remarkable and interesting range of problems. It has had considerable responsibility for the development of design computational procedures for air vehicles and airborne systems. As you are aware, the original REAC equipment was developed under SDC auspices.

The application of this equipment is dependent upon the analysis of many mathematical problems, such as the requirements of real time simulation, overall system analysis, and the use of digital check solutions. The Cyclone Laboratories of the Reeves Instrument Corporation pioneered the use of automatic digital computing equipment for check solutions for analog computation. The experience with the digital check solutions for analog equipment was later useful for developments such as the present project for the development of an operational flight trainer using digital computation.

In certain of these cases the actual problem in applied mathematics is more general than the specific solution of an individual system of differential

equations. The problems involve classes of systems. For instance, the development of analog computing equipment involves questions concerning ranges of behavior, and, similarly, the major objective in the development of an operational flight trainer using digital computation is a flexibility which will permit the consideration of a reasonable range of aircraft. Thus, it is characteristic of these problems that while they are not specifically tied to individual systems, nevertheless, like almost all problems in applied mathematics, they do have precise limits.

It seems to me that the first responsibility of an applied mathematician is to solve the given problem with precisely the generality given to him. If more than one solution is possible, his objective is to obtain the optimum solution. Until the given problem is solved, his major interest should not become involved in any generalization of the given problem. After a specific problem of a certain type has been solved, the solution itself may give some light on effective generalizations and tie-ins. The solution may be a small window which will permit us to see generalizations that are really valuable.

The work for SDC indicated above has led to two areas of development which I explored under ONR auspices with two colleagues. Dr. Miller and I have developed a general error theory for differential analyzers. (Cf. Ref. [3] and [7]) This is adequate for digital procedures as well as the more difficult continuous computer. J. Winson has modified this theory and applied the result to "linear" NEAC equipment with very interesting results. (Ref. [8]) Other developments of this theory were undertaken by the Advisory Board on Simulation of the University of Chicago. The latter will also publish an extended version of the theory. Dr. Brock and I have worked on certain numerical computing procedures which are applicable to the types of systems of differential equations which appear in problems for air vehicles. (Ref. [1] and [2])

At present, it seems to me that it is less expensive and more convenient to study a system of differential equations of the type which occurs in air vehicle problems by the use of a continuous differential analyzer properly supplemented by digital computation, than any other method. Parameters may be varied, equations "simplified", and the equivalent of major design changes carried out with ease. The person with engineering responsibility for the design is at all times in intimate contact with his problem. Furthermore, real time simulation at present seems to require the use of continuous equipment. I was pleased to note Mr. Wood in the present meeting has independently made similar remarks concerning the applicability of continuous computers.

There are, however, many problems associated with the use of such equipment. To check the initial setup and at the same time obtain a preliminary test of accuracy, a digital check problem is very useful. The size of these problems makes automatic computers necessary, and critical results can be checked by a digital computation.

While the use of digital check computation is probably the best manner of determining the accuracy of the results of a specific study, it is also desirable to obtain a general theory for the accuracy of continuous differential analyzers. Mathematically, this means that we must obtain a comparison between the solutions of two systems of differential equations which in some sense approximate each other.

Such a comparison is desirable and essential in many other problems besides error studies for differential analyzers. In almost every physical application of ordinary differential equations, the corresponding problem arises of comparing the solutions of a known system of differential equations with the solutions of another system of differential equations which approximates the first system.

Now if this problem involved only differential equations of the same order, the usual existence theory of ordinary differential equations which discusses the dependence on parameters of two systems of the same order would give an adequate answer. But both in the case of the error analysis and the other applications referred to, one has the situation that the two systems to be compared are of different order. Thus, mathematically, one has the problem of comparing two systems of differential equations

$$F_1 = 0$$

$$G_1 = 0$$

where the second system may be of higher order than the first, but the extra higher derivatives appear in the second system with relatively small partial derivatives. This is the problem which Dr. K. S. Miller and I solved in our joint paper in the M.I.T. Journal of Mathematics and Physics. (Ref. [3]) The review in Mathematical Reviews of this paper gives the incorrect impression that only a much more specialized problem was considered. Since the major objective of the paper was exactly this mathematical problem, it was quite disheartening to find that the generality of the result was ignored.

In order to tackle this problem associated with an increase in order, one must introduce a new technique. The basis of this technique is the following:

Around a given fixed solution, one may "linearize" the given system of equations, using the classical existence theory procedure. "Linearization" of course refers to a process of replacing the given system by a linear system in which one has constant coefficients. Now the usual rough method of attacking this is the following: One starts with a single solution given over a fixed interval of the independent variable t . One is trying to study the solutions with initial conditions close to this one. The fixed t interval is divided into shorter intervals and on each of these the given system F_1 is replaced by

a set of linear equations with constant coefficients and appropriate forcing terms. It is not uncommon to use this system to study the original, the claim then being that if one repeatedly subdivided the given interval and took initial conditions very near to those of the given solution, one would approximate the various other solutions.

It does not seem to be widely appreciated that the classical theory of differential equations permits one to "linearize" in a much more effective manner. We may have to use a subdivision of the t interval into a number of smaller intervals and use as the basic approximation a number of linear systems with constant coefficients, but one can base the discussion on a fixed subdivision of the t interval and the range of initial conditions can also be considered to be fixed and of finite extent.

Briefly, the process consists of introducing a system of equations,

$$F_1(\dots, \alpha) = 0,$$

with a parameter α with the following properties: The system,

$$F_1(\dots, 0) = 0,$$

is the basic approximation, i.e. it is a linear system with coefficients which are constant on intervals and change at most a finite number of times on the given interval. There is also a value α_0 for the parameter such that

$$F_1(\dots, \alpha_0) = 0$$

is the original, possibly non-linear, system that we wish to study.

We now consider the system

$$F_1(\dots, \alpha) = 0$$

This is a system of differential equations on unknown functions $x_1(t, \alpha)$ which normally can be considered to be analytic in the parameter α . A solution function $x_1(t, \alpha)$ then can be expanded in powers of α around $\alpha = 0$. Thus

$$x_1(t, \alpha) = x_1(t, 0) + \frac{\partial x_1}{\partial \alpha}(t, 0)\alpha + \frac{1}{2} \frac{\partial^2 x_1}{\partial \alpha^2}(t, 0)\alpha^2 + \dots$$

The setup must be such that this series converges for $\alpha = \alpha_0$. This is the single assumption made in the whole process and this question has to be explored in each individual case. However, the theory shows that even if one used a linearization with coefficients constant for the whole interval, there is a range of initial conditions for which one has convergence. It may, however, be more convenient and significant to linearize on subintervals.

The above discussion is based on the existence of a single "clothes line" solution along which the regions are pinned. K. S. Miller and I have written a book on Existence Theorems which will appear very shortly. (Cf. Ref. [4]) In this we indicate that there are essentially two classical situations for existence theorems, one of these is the "initial neighborhood" or "in the small" theory, the other requires the existence of such a running solution. In practice the existence of the running solution or a number of such can be established by computation, and thus computation and theory can be utilized to give a complete theory which is inaccessible by theory alone. Conversely, the theory gives precise local information not available from computation. This book is being published by the New York University Press.

Now the linearized equation gives the critical information required for the linearization. The $x_1(t, \alpha)$ satisfies the linearized system while the $\frac{\partial x_1}{\partial \alpha}$ satisfies a linear system of equations with the same homogeneous part but different forcing terms, and so do the $\frac{\partial^2 x_1}{\partial \alpha^2}$ and the higher partials.

Thus, our linearization process refers the original problem to the study of the $x_1, \frac{\partial x_1}{\partial \alpha}, \frac{\partial^2 x_1}{\partial \alpha^2}$. Each of these sets of functions satisfies a linear system of equations with constant or interval-constant coefficients. The systems for the different sets differ only in the forcing terms.

Now the above linearization permits us to attack our original problem of comparing the solutions of different systems by linearizing both the given

system and the approximating system. As far as basic stability is concerned, the variation in the forcing terms is insignificant. Thus, the final result of the linearization process is to set up a corresponding problem on a linear system of differential equations with constant coefficients or interval-constant coefficients.

This last problem can be directly attacked. In a linear system, the increase in order with small coefficients will introduce new modes with large modal values. Thus, the homogenous equations, for instance, will have solutions with terms in the form $e^{\lambda t}$ where $|\lambda|$ is large. If λ has a large positive real part, this term will make the entire solution unstable. If λ has a large negative real part, these terms disappear rapidly, the desired result. Furthermore, in this case by proper machine procedure the effects of the additional terms can be made negligible. There is an intermediate case in which rapid oscillations can appear.

These then are the major possible effects of the increase in order. We refer to these as the λ effect. In addition, a differential analyzer may be perturbed so as to change the equation without increasing the order. These we call α errors. A differential analyzer is also dependent on stored information. A perturbation in this stored information we call a β error. It is possible to analyze the growth of error in digital method of solution in terms of β errors. The truncation errors and the round-off effect at each step can be considered as a perturbation on stored information. (Cf. Ref. [1])

The use of digital check solutions requires a precise analysis of the error in the solution since the solution must be good enough to serve as a standard but not so good as to represent a waste of computation. By means of the above discussion of β errors, the problem can be referred back to the error at each step. In general this consists of a truncation error and round off errors.

Let us consider the truncation error. This will be dependent on the method of integration used. The earlier automatic sequence digital calculators had very limited memory available and it was urged that one use methods like the Runge-Kutta method which are not dependent on the storage of previously calculated results. However, this argument does not appeal to me except in the case of extremely limited storage and very simple equations since the total amount of computation is tripled in the more general cases. Instead I would prefer methods like the Milne "open" and "closed" or simply "open" procedures, where the previously calculated values of the derivatives are utilized over again to obtain higher order approximations. The additional storage required is seldom significant.

Among other advantages, one can readily calculate the truncation error for a p 'th order open process if the $(p + 2)$ nd derivative is reasonably constant over the interval involved in the integration process:

$$A_p h^{p+1} \frac{d^{(p+2)} x}{dt}.$$

One can obtain a similar expression for the truncation error for the closed step,

$$B_p h^{p+1} \frac{d^{(p+2)} x}{dt}$$

so that if one performs an open step and then a closed step, the difference is a fixed constant multiple of the truncation error for either step. One can show that there is relatively little advantage in going beyond the first closed step. (Cf. Ref. [5])

Since the error for the closed step is definitely smaller than that for the open procedure, this would indicate that the preferred method might be an open and closed process at each step. However, this means that one is doing

twice the work of a single open step, so that one should compare the open and closed procedure constant with an open procedure for one half the interval, i.e. E_p should be compared with $A_p/2^p$. If these quantities are comparable, the open method with the smaller interval is preferable since the hypotheses of the discussion are more strongly fulfilled.

I have used this procedure but with every tenth step an open followed by a closed so that one can estimate the errors. (Ref. [1])

The total error can be obtained in two ways. One is the detailed θ error method indicated above. On the other hand, the θ error theory shows that under reasonable general circumstances the total error is linear in the individual truncation errors. If possible, one should therefore repeat the procedure with half the interval. For if we use half the interval, the error is cut down to $1/2^p$ its former value so that the difference between the original solution and the solution with half the interval is close to the total error for the first. Incidentally, the "extrapolation to zero" process suggested by this is based on certain assumptions concerning the order of the error and in effect improves the answer only by an increase in one in the order. Such an improvement in the order could be obtained directly by the open integration process in the analytic case with less computation.

Normal numerical analysis is based on the approximation of arbitrary functions by polynomials. This is the reason for the assumption made above that the highest derivative should be essentially a constant. However, in the differential equations which we have considered, one can obtain an approximate expression valid over a larger region if, instead of using polynomials, one uses linear combinations of exponentials. Furthermore, the "linearization" referred to above seems to indicate that exponentials would be preferable and also what exponential functions to use.

This of course raised the problem of extending the usual results of numerical analysis to the situation in which linear combinations of exponentials are used instead of polynomials. One can regard the polynomial case, in fact, as a special case of the exponential.

This was basically the problem that Dr. Brock and I considered in our paper on the use of exponential sums in integration. (Cf. Ref. [2]) We considered specifically the processes for open and closed integration, and by the use of certain determinants we obtained formulae which described the truncation error for various types of integration. The method we used for evaluating the error is applicable to any integration process which uses equally spaced points and is linear in the value of the integrand. These of course include such methods as the open and closed Milne process, Runge-Kutta, etc. These formulae specialize in the case of polynomials to a result which is more convenient to use than the usual derivative relation.

BIBLIOGRAPHY

- [1] P. Brock and F. J. Murray, "Planning and error analysis for the numerical solution of a test system of differential equations on the IBM sequence calculator", Project Cyclone Report, Reeves Instrument Corporation, 215 E. 91 St., New York 28, N. Y. (1949)
- [2] P. Brock and F. J. Murray, "The use of exponential sums in step by step integration", Math. Tables and Other Aids to Computation, v. 6 (1952), p. 63-78 and p. 138-150.
- [3] K. S. Miller and F. J. Murray, "A mathematical basis for an error analysis of differential analyzers", Jour. of Math. and Physics, v. 32 (1953), p. 136-163.
- [4] K. S. Miller and F. J. Murray, "Existence theorems for ordinary differential equations", New York University Press, N. Y., N. Y. (1954).
- [5] F. J. Murray, "Planning and error considerations for the numerical solution of differential equations on a sequence calculator", Mathematical Tables and Other Aids to Computation, v. 4 (1950), p. 133-144.
- [6] F. J. Murray, "Error analysis for mathematical machines", Trans. of the N. Y. Acad. Sci., v. 13 (1951), p. 168-174.
- [7] F. J. Murray, "Mathematical error analysis for continuous computers", Project Cyclone Symposium II, Part II, Reeves Instrument Corporation, U. S. Navy Special Devices Center and U. S. Navy Bureau of Aeronautics, (1952) p. 139-146, New York.
- [8] J. Winson, "The error analysis of electronic analogue computation. Linear differential equations with constant coefficients", Ph. D. Thesis, Columbia University, (1954) N. Y., N. Y.